

ConsistencyChecker: Tree-based Evaluation of LLM Generalization Capabilities

Zhaochen Hong^{1*} Haofei Yu^{1*} Jiaxuan You¹

¹University of Illinois Urbana-Champaign

{zhong42,haofei2,jiaxuan}@illinois.edu

Abstract

Evaluating consistency in Large Language Models (LLMs) is crucial for ensuring reliability, particularly in complex, multi-step interactions between humans and LLMs. Traditional self-consistency methods often miss subtle semantic changes in natural language and functional shifts in code or equations, which can accumulate over multiple transformations. To address this, we propose *ConsistencyChecker*, a tree-based evaluation framework designed to measure consistency through sequences of reversible transformations, including machine translation tasks and AI-assisted programming tasks. In our proposed framework, nodes represent distinct text states, while edges correspond to pairs of inverse operations. Dynamic and LLM-generated benchmarks ensure a fair assessment of the model’s generalization ability and eliminate benchmark leakage. Consistency is quantified based on similarity across different depths of the transformation tree. Experiments on eight models from various families and sizes show that ConsistencyChecker can distinguish the performance of different models. Notably, our consistency scores, computed entirely without using WMT paired data, correlate strongly ($r > 0.7$) with WMT 2024 auto-ranking, demonstrating the validity of our benchmark-free approach. Our implementation is available at <https://github.com/ulab-uiuc/consistencychecker>.

1 Introduction

Large Language Models (LLMs) have emerged as transformative tools in artificial intelligence, demonstrating remarkable capabilities in natural language understanding (Minaee et al., 2024), generation (Minaee et al., 2024), and complex reasoning (Brown et al., 2020; Chowdhery et al., 2024). These models are increasingly deployed in applications such as multilingual translation pipelines (Xu

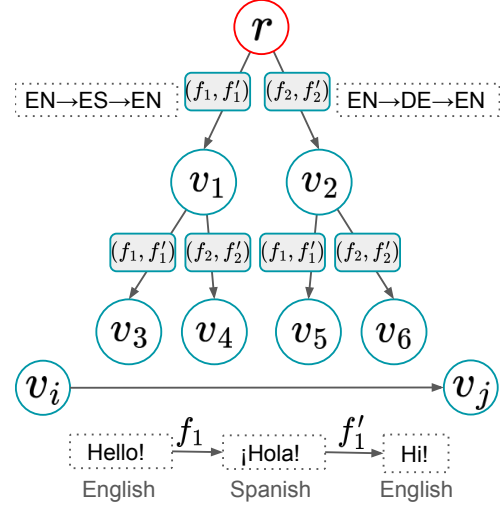


Figure 1: **Overview of the ConsistencyChecker.** It shows a self-consistency tree for evaluating LLMs on machine translations. The root node (r) is the initial English sentence generated by the *evaluator*. Subsequent nodes (v_i) are produced by the *evaluator* using transformation pairs, such as English→Spanish→English (f_1, f_1') and English→German→English (f_2, f_2'). The evaluation framework supports multilingual translations (e.g., French, Czech, Japanese) and can be extended to programming tasks.

et al., 2023) and automated code generation (Jiang et al., 2024b), all of which require maintaining semantic and functional consistency across multiple interactions. A fundamental research question thus arises: *Can we effectively evaluate the consistency of LLMs over multi-step transformations?*

Challenges for evaluating consistency. Evaluating LLM consistency across multi-step transformations is inherently difficult. Existing methods, such as self-consistency sampling (Min et al., 2023; Wang et al., 2022), which typically assess agreement between outputs for a single prompt, could fail to capture how meaning may drift over a sequence of transformations. This limitation is critical in tasks like machine translation and code generation, where minor changes at each iteration,

*Two authors contribute equally to this paper.

though locally coherent, can accumulate into major inconsistencies after multiple steps. For example, a translated sentence may gradually shift semantically, and a code snippet may become non-executable after multiple transformations. Effective evaluation must therefore assess *semantic* and *functional* consistency across the multi-step trajectory. For simplicity, we consider semantic consistency to be a subset of functional consistency, which mostly applies to natural language, whose functional consistency is equivalent to its semantic information-keeping consistency.

Challenges in evaluating generalization abilities.

Most benchmarks aim to evaluate how LLMs perform on unseen data in specific domains, which is to assess their generalization capabilities. Existing evaluation approaches primarily rely on dataset-centric benchmarks (e.g., HumanEval (Chen et al., 2021) for code generation). The dataset construction process includes human annotation or crawling of real-world data from the Internet. However, due to the vast and opaque nature of LLM training data, it has become increasingly difficult to determine whether benchmark examples were truly unseen during training. Several studies have found evidence of benchmark leakage (Xu et al., 2024; Zhou et al., 2025), where evaluation data overlaps with training corpora, undermining the reliability of performance metrics. Therefore, there is a growing need for novel benchmarking methodologies that dynamically generate evaluation data to ensure a more faithful assessment of model generalization.

Novel tree-based benchmark-free evaluation. To address these limitations, we introduce *ConsistencyChecker*, a *tree-based* and *benchmark-free* evaluation framework that operates without reference datasets, enabling cheap and domain-agnostic assessments of LLM generalization capabilities. At its core, *ConsistencyChecker* constructs self-consistency trees, where each node represents an LLM-generated state in transformations (e.g., output of one translation step), each edge denotes a pair of inverse operations, and each path in the self-consistency tree represents a multi-step transformation sequence operated by the LLM. For instance, starting with an English sentence, we perform a round-trip translation (English-French-English) executed by the LLM, yielding a slightly varied English sentence as a subsequent node. By analyzing consistency across nodes at different depths, *ConsistencyChecker* quantifies an LLM’s reliability in

retaining consistency (e.g. a sentence still conveys the same information, or code still gives the same output given the same inputs) through iterative transformations. *ConsistencyChecker* highlights two key innovations: (1) utilizing a tree structure to comprehensively evaluate consistency across multi-step transformation processes; (2) utilizing dynamically sampled data from LLMs to prevent benchmark leakage during evaluation.

Key findings. Based on our experimental results on both machine translation and AI-assisted programming tasks, we find that: (1) *ConsistencyChecker* can effectively distinguish the performance of different models and assign them varying scores, demonstrating its utility as a distinguishable metric for both tasks; (2) GPT-4o-mini ranks highest in machine translation tasks, while Qwen-2.5-32B performs best in code generation tasks, indicating that evaluation results align with these models’ performance on other benchmarks; (3) *ConsistencyChecker* can operate without relying on predefined benchmarks and still produces rankings highly correlated with expensive evaluations relying on paired data on the WMT 2024 benchmark.

2 Related Works

Round-trip consistency in translations. Machine translation evaluation has evolved from early rule-based systems (Weaver, 1952) to modern transformer-based approaches (Vaswani et al., 2017). Traditional metrics like BLEU (Papineni et al., 2002), ROUGE (Lin, 2004), and TER (Snover et al., 2006) rely on reference translations, while benchmarks like WMT 2024 (Barrault et al., 2019) depend on parallel corpora crawled from internet resources, often unavailable in low-resource languages or specialized domains. Round-trip translation, where text is translated between languages and back (van Zaanen and Zwarts, 2006), provides a benchmark-free quality checker analogous to our transformation chains. *ConsistencyChecker* generalizes this concept, extending it beyond translation to arbitrary functional-preserving operations (e.g., AI-assisted programming), thereby enabling evaluation in scenarios where parallel data or reference outputs are absent.

Consistency-based evaluation. Recent LLM evaluation frameworks, such as Divide-Conquer-Reasoning (Cui et al., 2024) and MT-Eval (Kwan et al., 2024), improve consistency assessment by breaking down tasks and using dynamic bench-

marks. However, they often miss tracking semantic or functional drift across complex, multi-step transformations (Wang et al., 2024). Our work builds on these ideas by introducing an explicit tree-structured evaluation, which systematically measures consistency at every step and path. Uniquely, we also use LLMs to generate dynamic benchmarks, enabling comprehensive and scalable consistency evaluation beyond existing methods.

Formal verification method. The development of reliable systems has traditionally depended on formal verification techniques such as model checking (Clarke et al., 1999), theorem proving (Nipkow et al., 2002), and SMT solvers (Barrett and Tinelli, 2018), all of which require significant computational resources and domain-specific specifications. The recent rise of LLMs for code generation has led to benchmarks like CodeXGLUE (Lu et al., 2021) and ProblemSolving (Hendrycks et al., 2021), but these approaches rely on curated datasets, which limits their general applicability. Inspired by both formal methods and the input/output-based evaluation used by online programming platforms like LeetCode, we instead use a set of test inputs to verify the functionality of LLM-edited Python code. Unlike traditional benchmarks, our approach does not require predefined correct outputs; instead, we concatenate and compare the outputs to assess functional consistency between different nodes.

3 Preliminary

In this section, we first provide detailed information on the definition of the self-consistency tree, which is the key concept for our proposed Consistency-Checker framework. Furthermore, we introduce two derived concepts of the self-consistency tree: paths and forests. These concepts are necessary for the calculation of our proposed consistency score.

3.1 Self-consistency Tree

To evaluate how well LLMs preserve consistency across reversible transformations, we begin by introducing our definition of operation, which refers to real-world tasks like translation or code editing defined by prompts and will be executed by the evaluatee LLM (the LLM to be evaluated). These operations are paired with their inverses to enable round-trip consistency checks. We then define nodes as intermediate model outputs and edges as paired operations. Together, they form the self-consistency tree, a structure that models the multi-

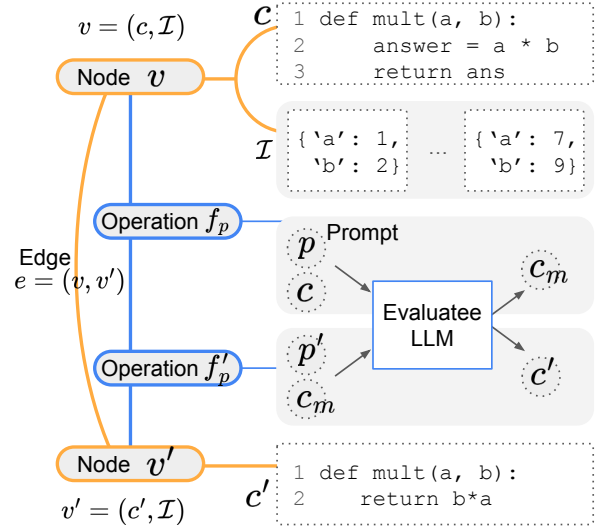


Figure 2: **Key concepts in the self-consistency tree (operation, node, and edge).** It provides a concrete example of the self-consistency tree in the AI-assisted programming task. The node v contains a function that returns the product of two positive integers, and \mathcal{I} is its set of inputs. The prompt p asks the evaluatee LLM to rewrite the code in node v to use a looped sum instead of just multiplying, while p' asks the same LLM to alter it back to simply multiplying.

step transformation process and allows consistency to be evaluated at different path lengths.

Operation (basic block). An operation is defined as a prompt-driven transformation performed by the evaluated LLM. For each prompt p , we define an operation f_p and its inverse f'_p , such that applying f_p followed by f'_p ideally recovers the original input. Formally,

$$f_p : c \rightarrow c', \quad (1)$$

where c is the input (e.g., code or text) and c' is the output after the transformation. The inverse operation f'_p is constructed so that $f'_p(f_p(c)) \approx c$ if the evaluated LLM is ideal.

Node (LLM-generated state). Each node $v = (c, \mathcal{I})$ is a tuple representing a single LLM generation and its associated test inputs: (1) $c \in \Sigma^*$ is the generated content (e.g., code, function, text); (2) $\mathcal{I} = [i_1, \dots, i_n]$ is a list of test inputs used to verify functional behavior. Here, Σ^* denotes the set of all possible strings. Regardless of output format (natural language, code, etc.), we treat each generated content c as an executable function for the evaluation of unified consistency. For semantic similarity evaluation in machine translation tasks, we consider it as a special case where \mathcal{I} is an empty set and the executed outcomes remain c .

Edge (paired operations). Each edge represents an operation and its inverse, $(f_p, f_{p'})$. It connects two nodes that are evaluated on the same set of test inputs, ensuring that functional consistency can be assessed under identical conditions. Formally, an edge $e_{ij} = (v_i, v_j)$ connects nodes $v_i = (c_i, \mathcal{I}_i)$ and $v_j = (c_j, \mathcal{I}_j)$ if and only if

$$\begin{cases} \mathcal{I}_i = \mathcal{I}_j, \\ c_j = f_{p'}(f_p(c_i)). \end{cases} \quad (2)$$

Self-consistency tree. Combining these concepts, a self-consistency tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ models the multi-step process by which an LLM transforms a root input through a sequence of operations. This hierarchical structure enables comprehensive measurement of how well the LLM preserves semantics or functionality over multiple transformation steps, with verification performed using the shared input set \mathcal{I} . Figure 1 illustrates the construction of nodes, edges, and operations. The tree generation process, detailed in Algorithm 1, uses operation pairs and root nodes. Notably, the branching factor at each layer of the tree is determined by the number of available operation pairs.

3.2 Derived Concepts: Path and Forest

Based on the definition of the self-consistency tree, we can extend it and define more useful concepts on top of it. These derived concepts help us calculate our proposed consistency score.

Path. A path in the self-consistency tree is defined as a sequence of nodes $P = (v_1, \dots, v_n)$, where each node is sequentially connected with the next through a valid edge in the self-consistency tree. The path represents a series of inverse operations applied by the evaluatee LLM, and its length corresponds to the number of operation pairs (the depth difference between the starting and ending nodes), allowing us to measure how consistency degrades across a certain number of transformation steps.

Forest. A forest $\mathcal{F} = \{\mathcal{T}_1, \dots, \mathcal{T}_M\}$ consists of multiple self-consistent trees, each rooted at a different initial node. These initial nodes are generated using distinct meta-prompts - e.g., "Write a 400-word paragraph in English about recent advances in computer science". This design allows us to analyze LLM consistency across diverse tasks and domains by comparing patterns across trees in the forest.

Algorithm 1 Tree Generation Algorithm

Require: Root node r , max depth D , operation pairs $((p_1, p'_1), \dots, (p_k, p'_k))$

Ensure: Tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$

```

1:  $\mathcal{V} \leftarrow \{r\}$ 
2:  $\mathcal{E} \leftarrow \emptyset$ 
3: for  $d \leftarrow 1$  to  $D$  do
4:    $U \leftarrow \{v \in \mathcal{V} \mid \mathcal{D}(v) = d - 1\}$ 
5:   for each  $v \in U$  do
6:     for  $i \leftarrow 1$  to  $k$  do
7:        $c' \leftarrow f_{p'_i}(f_{p_i}(v.c))$ 
8:        $v' \leftarrow (c', v.\mathcal{I})$ 
9:        $\mathcal{V} \leftarrow \mathcal{V} \cup \{v'\}$ 
10:       $\mathcal{E} \leftarrow \mathcal{E} \cup \{(v, v')\}$ 
11:     end for
12:   end for
13: end for
14: return  $(\mathcal{V}, \mathcal{E})$ 
```

4 ConsistencyChecker: Tree-based LLM evaluation for Generalization Ability

Based on the definitions in Section §3, we now detail the construction of consistency scores in ConsistencyChecker. We begin by defining the similarity score for node pairs, then extend this to path-level, tree-level, and finally forest-level scores, with each step building upon the previous one.

Node-pair similarity score. To quantify the functional and semantic similarity between two connected nodes $v_i = (c_i, \mathcal{I})$ and $v_j = (c_j, \mathcal{I})$ that share the same test inputs \mathcal{I} , we first obtain their execution outputs: $o_i = \text{exec}(c_i, \mathcal{I})$ and $o_j = \text{exec}(c_j, \mathcal{I})$. The similarity score $\text{sim}(v_i, v_j)$ is then computed between o_i and o_j , using measures such as the cosine similarity of semantic embeddings or the BLEU score.

Path-level consistency score. To evaluate whether a node still retains core functionalities after iterative transformations, we measure the end-to-end consistency between the initial and final nodes in a path. For a path on the tree $P = (v_1, \dots, v_n)$, the consistency score defined for that is:

$$C(P) := \text{sim}(v_1, v_n) \quad (3)$$

This metric evaluates whether the final output v_n remains functionally equivalent to the original v_1 after a sequence of transformations. For example, for code generation, applying operations of "add logging functionality" followed by "remove log-

ging functionality" should leave the core sorting functionality unchanged, as reflected by this score.

Tree-level consistency score. To assess how well an LLM preserves functional consistency over fixed-length transformation sequences, we define a tree-level consistency score with the help of the path-level consistency score. For a given depth n , we consider all valid n -step paths in the self-consistency tree as a set: $\mathcal{P}_n(\mathcal{T}) = \{(v_1, \dots, v_n) \mid (v_i, v_{i+1}) \in \mathcal{E} \text{ for } i = 1, \dots, n\}$. The tree-level consistency score is the average of path-level consistency scores over all such paths:

$$C_n(\mathcal{T}) := \frac{1}{|\mathcal{P}_n(\mathcal{T})|} \sum_{P \in \mathcal{P}_n(\mathcal{T})} C(P) \quad (4)$$

where $C(P)$ refers to the consistency score at the path level. This metric captures the model's ability to maintain core functionality across all n -step transformation sequences, reflecting both cumulative error and path-specific divergence.

Forest-level consistency score. To assess robustness and allow the measurement of generalization ability at specific domains, we define a consistency score over a forest of self-consistency trees $\mathcal{F} = \{\mathcal{T}_1, \dots, \mathcal{T}_M\}$. The score is computed as:

$$C_n(\mathcal{F}) := \frac{1}{M} \sum_{m=1}^M C_n(\mathcal{T}_m), \quad (5)$$

where $C_n(\mathcal{T})$ denotes the consistency at depth n of tree \mathcal{T} . Each tree is rooted at a node generated by the evaluator model. A forest with a high $C_1(\mathcal{F})$ score but a low $C_3(\mathcal{F})$ score indicates consistent performance for short transformation chains but degradation in longer transformation sequences.

ConsistencyChecker. In practice, setting a large value for n increases computational cost, while a small n leads to inaccurate estimates that overlook accumulated errors in multi-step transformations. To balance these trade-offs, we select $n = 3$ and use the forest-level consistency score $C_3(\mathcal{F})$ as the final metric in our *ConsistencyChecker* framework.

5 Evaluation Tasks and Meta Prompts

We evaluate ConsistencyChecker on two tasks: *machine translation* and *AI-assisted programming*. Each task is specified by a meta-prompt that generates a set of root nodes and paired operations for constructing self-consistency trees. In these cases, nodes correspond to either natural language sentences or executable code snippets. These tasks

Table 1: **Meta prompts for evaluation tasks.** Instructions about YAML and formatting are abbreviated.

Machine translation
Write a 400 word, complicated English paragraph that might appear on a news website. Please do this in a function way, e.g. provide a function called "main" that returns the content as a string.
AI-assisted programming
Write a LeetCode-Hard style problem. The problem must be super hard, even a graduate student in computer science will struggle to solve it. Do not attempt to generate long, nested dicts. But it will require a very long and complicated solution. The execution time should be very short. However, it does not need to be super long. It can be shorter, but it must be really hard. Please do this in a functional way, e.g., provide a function called "main" that returns the intended answer.

are representative, as they cover the full spectrum of node types—from static text with no external dependencies to code that serves as a solution to a LeetCode-style problem. The machine translation task focuses on assessing semantic consistency, while the AI-assisted programming task focuses on evaluating functional consistency. We set a time limit of 2 seconds per case in the test inputs to ensure that the verification process can end in a predictable time.

Machine translation. The root node of a translation task is a paragraph randomly sampled. Operations involve translating this paragraph into and from another language. We sample three languages from an evaluator, which are typically French, Spanish, and German. This setup ensures that the task captures the nuances of multilingual translation across diverse linguistic contexts.

AI-assisted programming. The root node for each programming task consists of a LeetCode-style coding problem and 20 associated test inputs. Operations are defined as transformations that prompt the model to implement the solution in a different, but functionally equivalent, way. This setup enables us to evaluate the model's ability to generate code that remains functionally and semantically consistent across varying levels of abstraction and complexity.

Connections between two tasks.. Although the translation and coding tasks differ in complexity and domain, both are implemented within the same framework. For translation, the function takes no inputs and simply returns the target natural language text. For the coding task, the function accepts inputs as required by the problem. This uni-

fied design makes our framework easily extensible to other tasks, such as step-by-step arithmetic, which can also be represented as Python code.

6 Experimental Settings

Evaluator models. Evaluator models determine the quality of dynamic benchmark generation. All experiments, except those in Table 4, use Qwen-2.5-72B as the evaluator model to generate dynamic benchmarks. For both tasks, we use the same evaluator-generated benchmarks across all models to ensure fair comparison and minimize the variability that fully on-the-fly benchmarks could introduce. Specifically, the evaluator model generates 10 root nodes for each task, which are then used consistently to evaluate all LLMs.

Evaluatee models. Using LLM-generated benchmarks, we evaluate a diverse set of language models across different families and sizes, including GPT-4o-mini (OpenAI et al., 2024), Qwen-2.5 (1.5B, 7B, 14B, 32B, 72B), and LLaMA-3.1 (8B, 70B) (Grattafiori et al., 2024), on both machine translation and AI-assisted programming tasks. Additionally, we assess the correlation between ConsistencyChecker and WMT 2024 metrics for machine translation benchmarks on specific language pairs. The models evaluated in this comparison include Claude-3.5-Sonnet (Anthropic, 2024), GPT-4, Gemini-1.5-Pro (Team et al., 2024), Mistral-Large (Jiang et al., 2023, 2024a), LLaMA-3-70B, and Phi-3-Medium (Abdin et al., 2024).

ConsistencyChecker settings. For the calculation of ConsistencyChecker, we fix $n = 3$, $M = 10$, and use embedding-based metrics such as NV-Embed-v2¹ to compute similarity scores between node pairs. We then calculate the forest-level consistency score as our main result. All reported scores are presented as means with standard deviations to reflect variability across evaluation runs.

7 Experimental Results

We show experimental results of 8 evaluated models on both machine translation and AI-assisted programming tasks, and analyze the trends and insights for the two tasks separately.

ConsistencyChecker is reliable for evaluating machine translation tasks. As shown in Table 2, larger models consistently achieve higher consistency scores than smaller models within the same

Table 2: **Evaluation results with ConsistencyChecker.** Each score represents LLMs’ ability to maintain semantic similarity in machine translation tasks and preserve functional consistency in programming tasks. Higher score indicates larger ratios for preserving the original information after transformations.

Model	Translation \uparrow	Programming \uparrow
GPT-4o-mini	98.0 ± 0.0	76.5 ± 2.7
Qwen-2.5-1.5B	80.3 ± 0.5	63.4 ± 1.4
Qwen-2.5-7B	90.0 ± 0.8	71.7 ± 0.4
Qwen-2.5-14B	94.7 ± 0.1	79.9 ± 1.0
Qwen-2.5-32B	96.4 ± 0.0	85.1 ± 1.1
Qwen-2.5-72B	97.2 ± 0.0	77.0 ± 1.9
LLaMA-3.1-8B	67.5 ± 3.0	60.4 ± 1.0
LLaMA-3.1-70B	71.9 ± 3.2	83.5 ± 1.0

Table 3: **Comparison between ConsistencyChecker and other metrics in WMT 2024 benchmarks.** ConsistencyChecker has a high correlation with reliable translation metrics like CometKiwi and AutoRank without utilizing paired data in WMT 2024.

Model	Ours \uparrow	CometKiwi \uparrow	AutoRank \downarrow
Czech–Ukrainian			
Claude-3.5-Sonnet	98.1	68.3	1.7
GPT-4	96.4	67.7	2.0
Gemini-1.5-Pro	97.5	66.8	2.0
Mistral-Large	96.5	66.6	2.3
LLaMA-3-70B	95.9	66.1	2.6
Phi-3-Medium	44.9	42.5	9.1
English–Chinese			
Claude-3.5-Sonnet	99.3	70.3	1.7
Gemini-1.5-Pro	99.3	69.8	1.8
GPT-4	98.1	69.3	2.0
Mistral-Large	97.5	66.5	2.8
LLaMA-3-70B	97.9	66.2	2.8
Phi-3-Medium	96.6	64.8	3.1

family for machine translation tasks. For example, Qwen-2.5-72B achieves a score of 97.2, which is 21.1% higher than Qwen-2.5-1.5B. Similarly, LLaMA-3.1-70B outperforms LLaMA-3.1-8B by 6.5%. Overall, models from the Qwen family generally achieve higher consistency scores than those from the LLaMA family, indicating that Qwen models produce more consistent results. Specially, we find GPT-4o-mini achieves the highest consistency score of 98.0, indicating the highest quality of it on machine translation tasks.

ConsistencyChecker is reliable for evaluating AI-assisted programming tasks. In AI-assisted programming tasks, larger models generally achieve higher functional consistency than smaller ones within the same model family, with Qwen-2.5-72B as an exception. Concretely, Qwen-2.5-32B scores 34.2% higher than Qwen-2.5-1.5B,

¹<https://huggingface.co/nvidia/NV-Embed-v2>

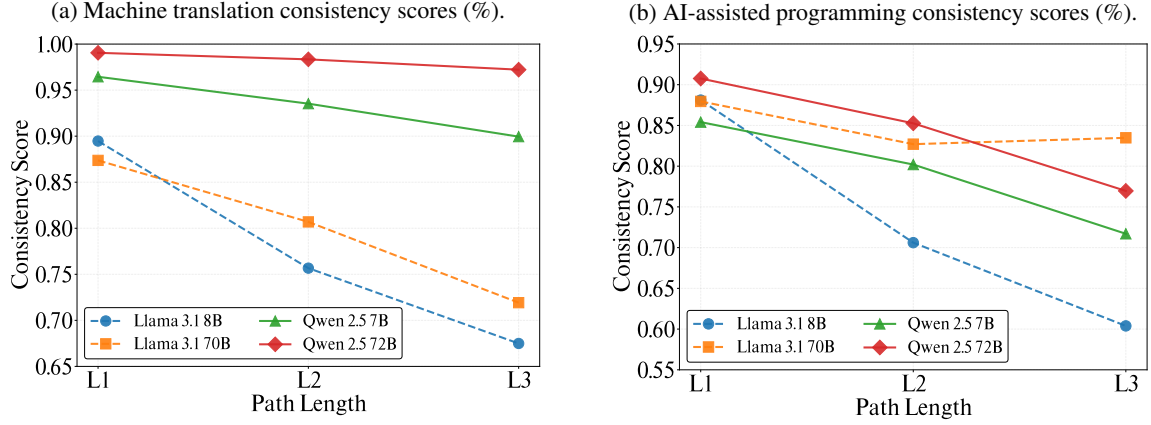


Figure 3: **Ablation study on path lengths.** For most models and for both translation and programming tasks, the consistency score becomes smaller when the path becomes longer.

while LLaMA-3.1-70B outperforms LLaMA-3.1-8B by 38.2%. These results underscore the importance of model scale for robust performance on complex transformation tasks.

ConsistencyChecker reaches high correlation with metrics in WMT 2024 without using its paired data. We compare ConsistencyChecker with the WMT 2024 (Kocmi et al., 2024) automatic rankings, focusing on two widely used metrics: CometKiwi (we refer to CometKiwi-DA-XL (Rei et al., 2023)) and AutoRank (the average of normalized CometKiwi and MetricX (Juraska et al., 2024) scores). Both metrics are computationally expensive, requiring large-scale pairwise human data or extensive model inference. In contrast, our approach is highly cost-efficient, as it does not require parallel data or data crawling—benchmarks are dynamically generated by LLMs. As shown in Table 3, we evaluate six representative LLMs—Claude-3.5-Sonnet, Gemini-1.5-Pro, GPT-4, LLaMA-3-70B, Mistral-Large, and Phi-3-Medium. For both CometKiwi and AutoRank, at least four out of six models are ranked similarly under ConsistencyChecker and WMT 2024. Pearson’s correlation coefficients, which are above 0.8 for both language pairs, provide strong evidence of the close agreement between ConsistencyChecker and established evaluation metrics. Detailed results are provided in Appendix §I.

8 Ablation Study

The path length n that we utilize for the calculation of ConsistencyChecker and the evaluator model that we rely on for dynamic benchmark generation are two key design elements that affect the performance of ConsistencyChecker. Therefore, we

conduct ablation studies on them.

Ablation on path length. Figure 3 shows that consistency scores generally decline as path length increases, reflecting greater transformation complexity. For example, in the machine translation task, the LLaMA-3.1-8B model shows a 24.6% drop in consistency from path length 1 to 3, while the Qwen-2.5-72B model shows a smaller 1.9% decrease. This trend holds across most models, regardless of family or size, though the rate of decline varies. In the AI-assisted programming task, the Qwen-2.5-7B model exhibits a 16.0% reduction, again confirming that errors tend to accumulate over longer transformation paths. Smaller drops in the consistency score indicate stronger self-consistency, as the model introduces fewer errors per operation. An exception occurs with the LLaMA-3.1-70B model in the programming task, where the consistency score is higher at path length 3 than at 2. This may be due to the model detecting and correcting syntax errors or bugs that appeared at intermediate steps.

Ablation on evaluator model. As shown in Table 4, different evaluator models (Qwen-2.5-7B and Qwen-2.5-72B) produce similar rankings across evaluated LLMs. For instance, in the translation task at path length 3, Qwen-2.5-7B assigns a score of 76.8, while Qwen-2.5-72B gives a slightly higher score of 78.2. In the programming task, for example, Qwen-2.5-7B evaluated by itself scores 72.2, while evaluation by Qwen-2.5-72B yields 71.7 (0.7% difference). We find that such an alignment becomes more pronounced at longer path lengths, where consistency patterns are clearer. Although absolute consistency scores may vary between evaluators, relative trends remain stable.

9 Discussion

In our main experiments, we use embedding-based models (NV-embed-v2) to compute node-pair similarity $\text{sim}(\cdot, \cdot)$ and calculate forest-level consistency score $C_3(\mathcal{F})$ based on that. However, BLEU (Papineni et al., 2002) remains a classical and widely adopted measure of translation quality. Therefore, we discuss their differences here.

BLEU is highly correlated but slightly worse than embedding-based models for machine translations. When evaluating translation across the five language pairs defined in WMT 2024, the BLEU-based consistency score is well correlated with WMT benchmark metrics, achieving Pearson correlations above 0.7 with three reference metrics in four out of five pairs. Furthermore, the BLEU-based $C_3(\mathcal{F})$ consistency score is also strongly correlated with the embedding-based $C_3(\mathcal{F})$, with average Pearson and Spearman correlations of 0.89 and 0.87, respectively, across all five pairs. However, embedding-based consistency scores are slightly more correlated with WMT benchmark metrics than BLEU-based scores when averaged across the five language pairs. Full evaluation results are provided in Appendix §I.

BLEU is surprisingly useful for evaluating programming tasks. For the AI-assisted programming task, the correlations were notably strong: Pearson coefficients reached 0.98 ($C_1(\mathcal{F})$), 0.97 ($C_2(\mathcal{F})$), and 0.99 ($C_3(\mathcal{F})$). This high correlation is unexpected, as BLEU is not designed to measure output functionality but is designed to measure n-gram overlaps. However, these results indicate that in this context, BLEU captures meaningful distinctions between outputs, likely because nodes are either fully functional (producing identical outputs for the same inputs), entirely different, or unexecutable due to code errors. As a result, even when outputs are not natural language, BLEU can effectively differentiate between them.

10 Case Study

In this section, we present two case studies involving different trees with the same root node but with different models for tree generation.

Leaf node from different models. The root node contains a function that returns an analysis of AI integration in healthcare systems. We construct self-consistency trees by applying a series of translation operations to the return value of the function:

Table 4: **Ablation study on evaluator models.** (Top) machine translation consistency scores (%) at complexity level $C_3(\mathcal{F})$ for model scales (7B–72B). (Bottom) AI-assisted programming consistency scores (%) at complexity level $C_3(\mathcal{F})$ for code transformations.

Evaluatee \ Evaluator	Qwen-2.5-7B	Qwen-2.5-72B
Machine translation		
Qwen-2.5-7B	76.8 \pm 5.1	90.0 \pm 0.8
Qwen-2.5-72B	78.2 \pm 7.3	97.2 \pm 0.0
LLaMA-3.1-8B	61.7 \pm 4.8	67.5 \pm 3.0
LLaMA-3.1-70B	30.8 \pm 9.2	71.9 \pm 3.2
AI-assisted programming		
Qwen-2.5-7B	72.2 \pm 1.9	71.7 \pm 0.4
Qwen-2.5-72B	79.4 \pm 2.3	77.0 \pm 1.9
LLaMA-3.1-8B	61.0 \pm 1.1	60.4 \pm 1.0
LLaMA-3.1-70B	81.5 \pm 2.1	83.5 \pm 1.0

Table 5: **Consistency score results when using BLEU to build ConsistencyChecker.** We select Qwen-2.5-72B as the evaluator and select different models as evaluatees. We utilize BLEU to build consistency scores.

Evaluatee	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$
Machine translation			
GPT-4o-mini	86.0 \pm 0.0	78.6 \pm 0.1	68.0 \pm 0.2
Qwen-2.5-1.5B	53.4 \pm 0.3	37.6 \pm 0.3	25.8 \pm 0.3
Qwen-2.5-7B	70.5 \pm 0.3	56.0 \pm 0.7	41.9 \pm 0.7
Qwen-2.5-14B	76.3 \pm 0.1	65.1 \pm 0.2	53.8 \pm 0.2
Qwen-2.5-32B	78.4 \pm 0.1	68.2 \pm 0.2	57.3 \pm 0.3
Qwen-2.5-72B	81.7 \pm 0.1	71.2 \pm 0.1	57.4 \pm 0.2
LLaMA-3.1-8B	59.5 \pm 1.2	47.6 \pm 1.1	36.6 \pm 0.9
LLaMA-3.1-70B	63.1 \pm 2.0	57.3 \pm 1.8	46.9 \pm 2.3
AI-assisted programming			
GPT-4o-mini	78.4 \pm 1.1	65.0 \pm 2.5	48.3 \pm 8.6
Qwen-2.5-1.5B	79.1 \pm 0.6	58.0 \pm 1.3	23.3 \pm 3.7
Qwen-2.5-7B	64.5 \pm 0.7	50.3 \pm 1.4	31.4 \pm 2.4
Qwen-2.5-14B	76.8 \pm 0.6	60.4 \pm 2.6	51.6 \pm 3.8
Qwen-2.5-32B	79.2 \pm 1.1	71.9 \pm 2.9	64.3 \pm 4.3
Qwen-2.5-72B	78.8 \pm 1.1	66.5 \pm 2.5	48.9 \pm 6.4
LLaMA-3.1-8B	69.4 \pm 1.1	42.9 \pm 2.4	11.7 \pm 2.1
LLaMA-3.1-70B	70.5 \pm 1.2	58.3 \pm 1.9	61.2 \pm 3.0

translating it into French, German, and Spanish, then back to English. These trees are evaluated with a path depth of 3, and we compare the outputs of LLaMA-3.1-8B and LLaMA-3.1-70B. The corresponding code for these leaf nodes is shown in Appendix §J. Based on the generated content of the leaf nodes, LLaMA-3.1-70B introduces fewer changes (Levenshtein distance: 66) compared to LLaMA-3.1-8B (Levenshtein distance: 215), indicating stronger consistency on larger models. This aligns with their consistency scores provided by ConsistencyChecker, highlighting the robustness

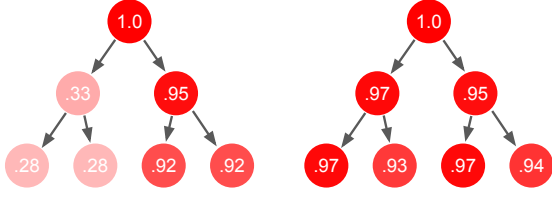


Figure 4: **Comparison of two self-consistency trees with different performances.** This diagram illustrates how Qwen-2.5-7B (left) and Qwen-2.5-72B (right) performed under the same root node and inverse operation pairs, in machine translation tasks. The diagram shows a simplified case where there are only 2 pairs of inverse operations. We use a lighter shade of red to indicate a smaller embedding similarity with the root node.

of larger models in maintaining meaning through multi-step transformations.

Features of under-performing self-consistency trees. We compare self-consistency trees generated by Qwen-2.5-7B and Qwen-2.5-72B from the same root node to analyze patterns of under-performance from the tree structure. The root consists of an around-290-token English paragraph discussing the impact of European Central Bank policies on global markets. Each model is prompted to translate the text to and from French, German, and Spanish. In the case of Qwen-2.5-7B, the French branch yields a significantly degraded output after round-trip translation, with only a 0.33 embedding similarity score to the root. The resulting English is just around 40 tokens and retains only partial information (see Figure 4). Subsequent nodes in this branch show equally low or worse similarity to the root node. By contrast, the German and Spanish branches maintain embedding similarities above 0.86. For Qwen-2.5-72B, all branches - including all child nodes - consistently achieve embedding similarities above 0.90. This illustrates a failure mode of less self-consistent models like Qwen-2.5-7B, and demonstrates how ConsistencyChecker effectively identifies and quantifies such inconsistencies with the help of the tree structure.

11 Conclusion

In this work, we propose ConsistencyChecker, an evaluation framework for assessing the generalization ability of LLMs effectively and efficiently. For our methods, we first introduce the concept of self-consistency trees, which model multi-step transformations as a tree structure. Based on this, we define a consistency score using a self-consistency

forest to enable robust evaluation across multiple tasks. Experimental results show that our consistency metric reliably estimates model performance for both translation and programming tasks, and closely matches WMT benchmark metrics such as AutoRank and CometKiwi for machine translation.

Limitations

While this study offers valuable insights into LLM consistency via the ConsistencyChecker framework, it has several limitations. The evaluation is limited to machine translation and code generation, excluding tasks like arithmetic. It primarily covers only 8 models in mainstream model families. Additionally, the study relies solely on automated metrics such as embedding similarity and BLEU, which may not fully reflect human judgment, especially for tasks requiring creativity or nuanced interpretation. Expanding task coverage, model diversity, and incorporating human evaluation in future work would strengthen the generalizability and robustness of the framework.

References

- Marah Abdin, Jyoti Aneja, Hany Awadalla, Ahmed Awadallah, Ammar Ahmad Awan, Nguyen Bach, Amit Bahree, Arash Bakhtiari, Jianmin Bao, Harkirat Behl, Alon Benhaim, Misha Bilenko, Johan Bjorck, Sébastien Bubeck, Martin Cai, Qin Cai, Vishrav Chaudhary, Dong Chen, Dongdong Chen, Weizhu Chen, Yen-Chun Chen, Yi-Ling Chen, Hao Cheng, Parul Chopra, Xiyang Dai, Matthew Dixon, Ronen Eldan, Victor Fragoso, Jianfeng Gao, Mei Gao, Min Gao, Amit Garg, Allie Del Giorno, Abhishek Goswami, Suriya Gunasekar, Emman Haider, Junheng Hao, Russell J. Hewett, Wenxiang Hu, Jamie Huynh, Dan Iter, Sam Ade Jacobs, Mojan Javaheripi, Xin Jin, Nikos Karampatziakis, Piero Kauffmann, Mahoud Khademi, Dongwoo Kim, Young Jin Kim, Lev Kurilenko, James R. Lee, Yin Tat Lee, Yuanzhi Li, Yunsheng Li, Chen Liang, Lars Liden, Xihui Lin, Zeqi Lin, Ce Liu, Liyuan Liu, Mengchen Liu, Weishung Liu, Xiaodong Liu, Chong Luo, Piyush Madan, Ali Mahmoudzadeh, David Majercak, Matt Mazzola, Caio César Teodoro Mendes, Arindam Mitra, Hardik Modi, Anh Nguyen, Brandon Norick, Barun Patra, Daniel Perez-Becker, Thomas Portet, Reid Pryzant, Heyang Qin, Marko Radmilac, Liliang Ren, Gustavo de Rosa, Corby Rosset, Sambudha Roy, Olatunji Ruwase, Olli Saarikivi, Amin Saied, Adil Salim, Michael Santacrose, Shital Shah, Ning Shang, Hiteshi Sharma, Yelong Shen, Swadheen Shukla, Xia Song, Masahiro Tanaka, Andrea Tupini, Praneetha Vaddamanu, Chunyu Wang, Guanhua Wang, Lijuan Wang, Shuohang Wang, Xin Wang, Yu Wang, Rachel Ward, Wen Wen, Philipp Witte, Haiping Wu, Xiaoxia Wu, Michael Wyatt, Bin Xiao, Can Xu, Jiahang Xu,

- Weijian Xu, Jilong Xue, Sonali Yadav, Fan Yang, Jianwei Yang, Yifan Yang, Ziyi Yang, Donghan Yu, Lu Yuan, Chenruidong Zhang, Cyril Zhang, Jianwen Zhang, Li Lyna Zhang, Yi Zhang, Yue Zhang, Yunan Zhang, and Xiren Zhou. 2024. [Phi-3 technical report: A highly capable language model locally on your phone](#). *Preprint*, arXiv:2404.14219.
- Anthropic. 2024. [The claude 3 model family: Opus, sonnet, haiku](#). In *Anthropic Model Card*.
- Loïc Barrault, Ondřej Bojar, Marta R. Costa-jussà, Christian Federmann, Mark Fishel, Yvette Graham, Barry Haddow, Matthias Huck, Philipp Koehn, Shervin Malmasi, Christof Monz, Mathias Müller, Santanu Pal, Matt Post, and Marcos Zampieri. 2019. [Findings of the 2019 conference on machine translation \(WMT19\)](#). In *Proceedings of the Fourth Conference on Machine Translation (Volume 2: Shared Task Papers, Day 1)*, pages 1–61, Florence, Italy. Association for Computational Linguistics.
- Clark Barrett and Cesare Tinelli. 2018. Satisfiability modulo theories. *Handbook of model checking*, pages 305–343.
- Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. [Language models are few-shot learners](#). *CoRR*, abs/2005.14165.
- Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde De Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, et al. 2021. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*.
- Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayanan Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. 2024. [Palm: scaling language modeling with pathways](#). *J. Mach. Learn. Res.*, 24(1).
- Edmund M Clarke, Orna Grumberg, and Doron Peled. 1999. *Model checking*. MIT press.
- Wendi Cui, Zhuohang Li, Damien Lopez, Kamalika Das, Bradley A. Malin, Sricharan Kumar, and Jiaxin Zhang. 2024. [Divide-conquer-reasoning for consistency evaluation and automatic improvement of large language models](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 334–361, Miami, Florida, US. Association for Computational Linguistics.
- Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.
- Dan Hendrycks, Steven Basart, Saurav Kadavath, Mantas Mazeika, Akul Arora, Ethan Guo, Collin Burns, Samir Puranik, Horace He, Dawn Song, and Jacob Steinhardt. 2021. [Measuring coding challenge competence with APPS](#). In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*.
- Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, L  lio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timoth  e Lacroix, and William El Sayed. 2023. [Mistral 7b](#). *Preprint*, arXiv:2310.06825.
- Albert Q. Jiang, Alexandre Sablayrolles, Antoine Roux, Arthur Mensch, Blanche Savary, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Emma Bou Hanna, Florian Bressand, Gianna Lengyel, Guillaume Bour, Guillaume Lample, L  lio Renard Lavaud, Lucile Saulnier, Marie-Anne Lachaux, Pierre Stock, Sandeep Subramanian, Sophia Yang, Szymon Antoniak, Teven Le Scao, Th  ophile Gervet, Thibaut Lavril, Thomas Wang, Timoth  e Lacroix, and William El Sayed. 2024a. [Mixtral of experts](#). *Preprint*, arXiv:2401.04088.
- Juyong Jiang, Fan Wang, Jiasi Shen, Sungju Kim, and Sunghun Kim. 2024b. A survey on large language models for code generation. *arXiv preprint arXiv:2406.00515*.
- Juraj Juraska, Daniel Deutsch, Mara Finkelstein, and Markus Freitag. 2024. Metricx-24: The google submission to the wmt 2024 metrics shared task. *arXiv preprint arXiv:2410.03983*.
- Tom Kocmi, Eleftherios Avramidis, Rachel Bawden, Ondrej Bojar, Anton Dvorkovich, Christian Federmann, Mark Fishel, Markus Freitag, Thamme Gowda,

- Roman Grundkiewicz, et al. 2024. Preliminary wmt24 ranking of general mt systems and llms. *arXiv preprint arXiv:2407.19884*.
- Wai-Chung Kwan, Xingshan Zeng, Yuxin Jiang, Yufei Wang, Liangyou Li, Lifeng Shang, Xin Jiang, Qun Liu, and Kam-Fai Wong. 2024. *Mt-eval: A multi-turn capabilities evaluation benchmark for large language models*. *Preprint*, arXiv:2401.16745.
- Chin-Yew Lin. 2004. *ROUGE: A package for automatic evaluation of summaries*. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.
- Shuai Lu, Daya Guo, Shuo Ren, Junjie Huang, Alexey Svyatkovskiy, Ambrosio Blanco, Colin Clement, Dawn Drain, Daxin Jiang, Duyu Tang, Ge Li, Lidong Zhou, Linjun Shou, Long Zhou, Michele Tufano, MING GONG, Ming Zhou, Nan Duan, Neel Sundaresan, Shao Kun Deng, Shengyu Fu, and Shujie LIU. 2021. *CodeXGLUE: A machine learning benchmark dataset for code understanding and generation*. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1)*.
- Marcus J Min, Yangruibo Ding, Luca Buratti, Saurabh Pujar, Gail Kaiser, Suman Jana, and Baishakhi Ray. 2023. Beyond accuracy: Evaluating self-consistency of code large language models with identitychain. *arXiv preprint arXiv:2310.14053*.
- Shervin Minaee, Tomas Mikolov, Narjes Nikzad, Meysam Chenaghlu, Richard Socher, Xavier Amatriain, and Jianfeng Gao. 2024. *Large language models: A survey*. *Preprint*, arXiv:2402.06196.
- Tobias Nipkow, Markus Wenzel, and Lawrence C. Paulson. 2002. *Isabelle/HOL: a proof assistant for higher-order logic*. Springer-Verlag, Berlin, Heidelberg.
- OpenAI, :, Aaron Hurst, Adam Lerer, Adam P. Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, Aleksander Madry, Alex Baker-Whitcomb, Alex Beutel, Alex Borzunov, Alex Carney, Alex Chow, Alex Kirillov, Alex Nichol, Alex Paino, Alex Renzin, Alex Tachard Passos, Alexander Kirillov, Alexi Christakis, Alexis Conneau, Ali Kamali, Allan Jabri, Allison Moyer, Allison Tam, Amadou Crookes, Amin Tootoochian, Amin Tootoonchian, Ananya Kumar, Andrea Vallone, Andrej Karpathy, Andrew Braunstein, Andrew Cann, Andrew Codisoti, Andrew Galu, Andrew Kondrich, Andrew Tulloch, Andrew Mishchenko, Angela Baek, Angela Jiang, Antoine Pelisse, Antonia Woodford, Anuj Gosalia, Arka Dhar, Ashley Pantuliano, Avi Nayak, Avital Oliver, Barret Zoph, Behrooz Ghorbani, Ben Leimberger, Ben Rossen, Ben Sokolowsky, Ben Wang, Benjamin Zweig, Beth Hoover, Blake Samic, Bob McGrew, Bobby Spero, Bogo Gierltler, Bowen Cheng, Brad Lightcap, Brandon Walkin, Brendan Quinn, Brian Guarraci, Brian Hsu, Bright Kellogg, Brydon Eastman, Camillo Lugaresi, Carroll Wainwright, Cary Bassin, Cary Hudson, Casey Chu, Chad Nelson, Chak Li, Chan Jun Shern, Channing Conger, Charlotte Barette, Chelsea Voss, Chen Ding, Cheng Lu, Chong Zhang, Chris Beaumont, Chris Hallacy, Chris Koch, Christian Gibson, Christina Kim, Christine Choi, Christine McLeavey, Christopher Hesse, Claudia Fischer, Clemens Winter, Coley Czarnecki, Colin Jarvis, Colin Wei, Constantin Koumouzelis, Dane Sherburn, Daniel Kappler, Daniel Levin, Daniel Levy, David Carr, David Farhi, David Mely, David Robinson, David Sasaki, Denny Jin, Dev Valladares, Dimitris Tsipras, Doug Li, Duc Phong Nguyen, Duncan Findlay, Edede Oiwoh, Edmund Wong, Ehsan Asdar, Elizabeth Proehl, Elizabeth Yang, Eric Antonow, Eric Kramer, Eric Peterson, Eric Sigler, Eric Wallace, Eugene Brevdo, Evan Mays, Farzad Khorasani, Felipe Petroski Such, Filippo Raso, Francis Zhang, Fred von Lohmann, Freddie Sulit, Gabriel Goh, Gene Oden, Geoff Salmon, Giulio Starace, Greg Brockman, Hadi Salman, Haiming Bao, Haitang Hu, Hannah Wong, Haoyu Wang, Heather Schmidt, Heather Whitney, Heewoo Jun, Hendrik Kirchner, Henrique Ponde de Oliveira Pinto, Hongyu Ren, Huiwen Chang, Hyung Won Chung, Ian Kivlichan, Ian O’Connell, Ian O’Connell, Ian Osband, Ian Silber, Ian Sohl, Ibrahim Okuyucu, Ikai Lan, Ilya Kostrikov, Ilya Sutskever, Ingmar Kanitscheider, Ishaan Gulrajani, Jacob Coxon, Jacob Menick, Jakub Pachocki, James Aung, James Betker, James Crooks, James Lennon, Jamie Kiros, Jan Leike, Jane Park, Jason Kwon, Jason Phang, Jason Teplitz, Jason Wei, Jason Wolfe, Jay Chen, Jeff Harris, Jenia Vavva, Jessica Gan Lee, Jessica Shieh, Ji Lin, Jiahui Yu, Jiayi Weng, Jie Tang, Jieqi Yu, Joanne Jang, Joaquin Quinonero Candela, Joe Beutler, Joe Landers, Joel Parish, Johannes Heidecke, John Schulman, Jonathan Lachman, Jonathan McKay, Jonathan Uesato, Jonathan Ward, Jong Wook Kim, Joost Huizinga, Jordan Sitkin, Jos Kraaijeveld, Josh Gross, Josh Kaplan, Josh Snyder, Joshua Achiam, Joy Jiao, Joyce Lee, Juntang Zhuang, Justyn Harriman, Kai Fricke, Kai Hayashi, Karan Singhal, Katy Shi, Kevin Karthik, Kayla Wood, Kendra Rimbach, Kenny Hsu, Kenny Nguyen, Keren Gu-Lemberg, Kevin Button, Kevin Liu, Kiel Howe, Krithika Muthukumar, Kyle Luther, Lama Ahmad, Larry Kai, Lauren Itow, Lauren Workman, Leher Pathak, Leo Chen, Li Jing, Lia Guy, Liam Fedus, Liang Zhou, Lien Mamitsuka, Lillian Weng, Lindsay McCallum, Lindsey Held, Long Ouyang, Louis Feuvrier, Lu Zhang, Lukas Kondraciuk, Lukasz Kaiser, Luke Hewitt, Luke Metz, Lyric Doshi, Mada Aflak, Maddie Simens, Madelaine Boyd, Madeleine Thompson, Marat Dukhan, Mark Chen, Mark Gray, Mark Hudnall, Marvin Zhang, Marwan Aljubei, Mateusz Litwin, Matthew Zeng, Max Johnson, Maya Shetty, Mayank Gupta, Meghan Shah, Mehmet Yatbaz, Meng Jia Yang, Mengchao Zhong, Mia Glaese, Mianna Chen, Michael Janner, Michael Lampe, Michael Petrov, Michael Wu, Michele Wang, Michelle Fradin, Michelle Pokrass, Miguel Castro, Miguel Oom Temudo de Castro, Mikhail Pavlov, Miles Brundage, Miles Wang, Minal Khan, Mira Murati, Mo Bavarian, Molly Lin, Murat Yesildal, Nacho Soto, Natalia Gimelshein, Na-

- talie Cone, Natalie Staudacher, Natalie Summers, Natan LaFontaine, Neil Chowdhury, Nick Ryder, Nick Stathas, Nick Turley, Nik Tezak, Niko Felix, Nithanth Kudige, Nitish Keskar, Noah Deutsch, Noel Bundick, Nora Puckett, Ofir Nachum, Ola Okelola, Oleg Boiko, Oleg Murk, Oliver Jaffe, Olivia Watkins, Olivier Godement, Owen Campbell-Moore, Patrick Chao, Paul McMillan, Pavel Belov, Peng Su, Peter Bak, Peter Bakkum, Peter Deng, Peter Dolan, Peter Hoeschele, Peter Welinder, Phil Tillet, Philip Pronin, Philippe Tillet, Prafulla Dhariwal, Qiming Yuan, Rachel Dias, Rachel Lim, Rahul Arora, Rajan Troll, Randall Lin, Rapha Gontijo Lopes, Raul Puri, Reah Miyara, Reimar Leike, Renaud Gaubert, Reza Zamani, Ricky Wang, Rob Donnelly, Rob Honsby, Rocky Smith, Rohan Sahai, Rohit Ramchandani, Romain Huet, Rory Carmichael, Rowan Zellers, Roy Chen, Ruby Chen, Ruslan Nigmatullin, Ryan Cheu, Saachi Jain, Sam Altman, Sam Schoenholz, Sam Toizer, Samuel Miserendino, Sandhini Agarwal, Sara Culver, Scott Ethersmith, Scott Gray, Sean Grove, Sean Metzger, Shamez Hermani, Shantanu Jain, Shengjia Zhao, Sherwin Wu, Shino Jomoto, Shiron Wu, Shuaiqi, Xia, Sonia Phene, Spencer Papay, Srinivas Narayanan, Steve Coffey, Steve Lee, Stewart Hall, Suchir Balaji, Tal Broda, Tal Stramer, Tao Xu, Tarun Gogineni, Taya Christianson, Ted Sanders, Tejal Patwardhan, Thomas Cunningham, Thomas Degry, Thomas Dimson, Thomas Raoux, Thomas Shadwell, Tianhao Zheng, Todd Underwood, Todor Markov, Toki Sherbakov, Tom Rubin, Tom Stasi, Tomer Kaftan, Tristan Heywood, Troy Peterson, Tyce Walters, Tyna Eloundou, Valerie Qi, Veit Moeller, Vinnie Monaco, Vishal Kuo, Vlad Fomenko, Wayne Chang, Weiye Zheng, Wenda Zhou, Wesam Manassra, Will Sheu, Wojciech Zaremba, Yash Patil, Yilei Qian, Yongjik Kim, Youlong Cheng, Yu Zhang, Yuchen He, Yuchen Zhang, Yujia Jin, Yunxing Dai, and Yury Malkov. 2024. [Gpt-4o system card](#). *Preprint*, arXiv:2410.21276.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. [Bleu: a method for automatic evaluation of machine translation](#). In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.
- Ricardo Rei, Nuno M. Guerreiro, José Pombal, Daan van Stigt, Marcos Treviso, Luisa Coheur, José G. C. de Souza, and André F. T. Martins. 2023. [Scaling up cometkiwi: Unbabel-ist 2023 submission for the quality estimation shared task](#). *Preprint*, arXiv:2309.11925.
- Matthew Snover, Bonnie Dorr, Rich Schwartz, Linnea Micciulla, and John Makhoul. 2006. [A study of translation edit rate with targeted human annotation](#). In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*, pages 223–231, Cambridge, Massachusetts, USA. Association for Machine Translation in the Americas.
- Gemini Team, Petko Georgiev, Ving Ian Lei, Ryan Burnell, Libin Bai, Anmol Gulati, Garrett Tanzer, Damien Vincent, Zhufeng Pan, Shibo Wang, Soroosh Mariooryad, Yifan Ding, Xinyang Geng, Fred Alcober, Roy Frostig, Mark Omernick, Lexi Walker, Cosmin Paduraru, Christina Sorokin, Andrea Tacchetti, Colin Gaffney, Samira Daruki, Olcan Serincoglu, Zach Gleicher, Juliette Love, Paul Voigtlaender, Rohan Jain, Gabriela Surita, Kareem Mohamed, Rory Blevins, Junwhan Ahn, Tao Zhu, Kornaphop Kawintiranon, Orhan Firat, Yiming Gu, Yujing Zhang, Matthew Rahtz, Manaal Faruqui, Natalie Clay, Justin Gilmer, JD Co-Reyes, Ivo Penchev, Rui Zhu, Nobuyuki Morioka, Kevin Hui, Krishna Haridasan, Victor Campos, Mahdis Mahdih, Mandy Guo, Samer Hassan, Kevin Kilgour, Arpi Vezzer, Heng-Tze Cheng, Raoul de Liedekerke, Siddharth Goyal, Paul Barham, DJ Strouse, Seb Noury, Jonas Adler, Mukund Sundararajan, Sharad Vikram, Dmitry Lepikhin, Michela Paganini, Xavier Garcia, Fan Yang, Dasha Valter, Maja Trebacz, Kiran Vodrahalli, Chulayuth Asawaroengchai, Roman Ring, Norbert Kalb, Livio Baldini Soares, Siddhartha Brahma, David Steiner, Tianhe Yu, Fabian Mentzer, Antoine He, Lucas Gonzalez, Bibo Xu, Raphael Lopez Kaufman, Laurent El Shafey, Junhyuk Oh, Tom Hennigan, George van den Driessche, Seth Odoom, Mario Lucic, Becca Roelofs, Sid Lall, Amit Marathe, Betty Chan, Santiago Ontanon, Luheng He, Denis Teplyashin, Jonathan Lai, Phil Crone, Bogdan Damoc, Lewis Ho, Sebastian Riedel, Karel Lenc, Chih-Kuan Yeh, Aakanksha Chowdhery, Yang Xu, Mehran Kazemi, Ehsan Amid, Anastasia Petrushkina, Kevin Swersky, Ali Khodaei, Gowoon Chen, Chris Larkin, Mario Pinto, Geng Yan, Adria Puigdomenech Badia, Piyush Patil, Steven Hansen, Dave Orr, Sebastien M. R. Arnold, Jordan Grimstad, Andrew Dai, Sholto Douglas, Rishika Sinha, Vikas Yadav, Xi Chen, Elena Gribovskaya, Jacob Austin, Jeffrey Zhao, Kaushal Patel, Paul Komarek, Sophia Austin, Sebastian Borgeaud, Linda Friso, Abhimanyu Goyal, Ben Caine, Kris Cao, Da-Woon Chung, Matthew Lamm, Gabe Barth-Maron, Thais Kagohara, Kate Olszewska, Mia Chen, Kaushik Shivakumar, Rishabh Agarwal, Harshal Godhia, Ravi Rajwar, Javier Snider, Xerxes Dotiwalla, Yuan Liu, Aditya Barua, Victor Ungureanu, Yuan Zhang, Bat-Orgil Batsaikhan, Mateo Wirth, James Qin, Ivo Danihelka, Tulsee Doshi, Martin Chadwick, Jilin Chen, Sanil Jain, Quoc Le, Arjun Kar, Madhu Gurusurthy, Cheng Li, Ruoxin Sang, Fangyu Liu, Lampros Lamprou, Rich Munoz, Nathan Lintz, Harsh Mehta, Heidi Howard, Malcolm Reynolds, Lora Aroyo, Quan Wang, Lorenzo Blanco, Albin Cassirer, Jordan Griffith, Dipanjan Das, Stephan Lee, Jakub Sygnowski, Zach Fisher, James Besley, Richard Powell, Zafarali Ahmed, Dominik Paulus, David Reitter, Zalan Borsos, Rishabh Joshi, Aedan Pope, Steven Hand, Vittorio Selo, Vihan Jain, Nikhil Sethi, Megha Goel, Takaki Makino, Rhys May, Zhen Yang, Johan Schalkwyk, Christina Butterfield, Anja Hauth, Alex Goldin, Will Hawkins, Evan Senter, Sergey Brin, Oliver Woodman, Marvin Ritter, Eric Noland, Minh Giang, Vijay Bolina, Lisa Lee, Tim Blyth, Ian Mackinnon, Machel Reid,

Obaid Sarvana, David Silver, Alexander Chen, Lily Wang, Loren Maggiore, Oscar Chang, Nithya Ataluri, Gregory Thornton, Chung-Cheng Chiu, Oskar Bunyan, Nir Levine, Timothy Chung, Evgenii Eltyshv, Xiance Si, Timothy Lillicrap, Demetra Brady, Vaibhav Aggarwal, Boxi Wu, Yuanzhong Xu, Ross McIlroy, Kartikeya Badola, Paramjit Sandhu, Erica Moreira, Wojciech Stokowiec, Ross Hemley, Dong Li, Alex Tudor, Pranav Shyam, Elahe Rahimtoroghi, Salem Haykal, Pablo Sprechmann, Xiang Zhou, Diana Mincu, Yujia Li, Ravi Addanki, Kalpesh Krishna, Xiao Wu, Alexandre Frechette, Matan Eyal, Allan Dafoe, Dave Lacey, Jay Whang, Thi Avrahami, Ye Zhang, Emanuel Taropa, Hanzhao Lin, Daniel Toyama, Eliza Rutherford, Motoki Sano, HyunJeong Choe, Alex Tomala, Chalence Safranek-Shrader, Nora Kassner, Mantas Pajarskas, Matt Harvey, Sean Sechrist, Meire Fortunato, Christina Lyu, Gamaleldin Elsayed, Chenkai Kuang, James Lottes, Eric Chu, Chao Jia, Chih-Wei Chen, Peter Humphreys, Kate Baumli, Connie Tao, Rajkumar Samuel, Cicero Nogueira dos Santos, Anders Andreassen, Nemanja Rakićević, Dominik Grewe, Aviral Kumar, Stephanie Winkler, Jonathan Caton, Andrew Brock, Sid Dalmia, Hannah Sheahan, Iain Barr, Yingjie Miao, Paul Natsev, Jacob Devlin, Feryal Behbahani, Flavien Prost, Yanhua Sun, Artiom Myaskovsky, Thanumalayan Sankaranarayanan Pillai, Dan Hurt, Angeliki Lazaridou, Xi Xiong, Ce Zheng, Fabio Pardo, Xiaowei Li, Dan Horgan, Joe Stanton, Moran Ambar, Fei Xia, Alejandro Lince, Mingqiu Wang, Basil Mustafa, Albert Webson, Hyo Lee, Rohan Anil, Martin Wicke, Timothy Dozat, Abhishek Sinha, Enrique Piqueras, Elahe Dabir, Shyam Upadhyay, Anudhyan Boral, Lisa Anne Hendricks, Corey Fry, Josip Djolonga, Yi Su, Jake Walker, Jane Labanowski, Ronny Huang, Vedant Misra, Jeremy Chen, RJ Skerry-Ryan, Avi Singh, Shruti Rijhwani, Dian Yu, Alex Castro-Ros, Beer Changpinyo, Romina Datta, Sumit Bagri, Arnar Mar Hrafnkels-son, Marcello Maggioni, Daniel Zheng, Yury Sulsky, Shaobo Hou, Tom Le Paine, Antoine Yang, Jason Riesa, Dominika Rogozinska, Dror Marcus, Dalia El Badawy, Qiao Zhang, Luyu Wang, Helen Miller, Jeremy Greer, Lars Lowe Sjos, Azade Nova, Heiga Zen, Rahma Chaabouni, Mihaela Rosca, Jiepu Jiang, Charlie Chen, Ruibo Liu, Tara Sainath, Maxim Krikun, Alex Polozov, Jean-Baptiste Lepiau, Josh Newlan, Zeyncep Cankara, Soo Kwak, Yunhan Xu, Phil Chen, Andy Coenen, Clemens Meyer, Katerina Tsihlias, Ada Ma, Juraj Gottweis, Jinwei Xing, Chenjie Gu, Jin Miao, Christian Frank, Zeynep Cankara, Sanjay Ganapathy, Ishita Dasgupta, Steph Hughes-Fitt, Heng Chen, David Reid, Keran Rong, Hongmin Fan, Joost van Amersfoort, Vincent Zhuang, Aaron Cohen, Shixiang Shane Gu, Anhad Mohanane, Anastasija Ilic, Taylor Tobin, John Wieting, Anna Bortsova, Phoebe Thacker, Emma Wang, Emily Caveness, Justin Chiu, Eren Sezener, Alex Kaskasoli, Steven Baker, Katie Millican, Mohamed Elhawaty, Kostas Aisopos, Carl Lebsack, Nathan Byrd, Hanjun Dai, Wenhao Jia, Matthew Wiethoff, Elnaz Davoodi, Albert Weston, Lakshman Yagati, Arun Ahuja, Isabel Gao, Golan Pundak, Susan Zhang, Michael Azzam,

Khe Chai Sim, Sergi Caelles, James Keeling, Abhanshu Sharma, Andy Swing, YaGuang Li, Chenxi Liu, Carrie Grimes Bostock, Yamini Bansal, Zachary Nado, Ankesh Anand, Josh Lipschultz, Abhijit Karmarkar, Lev Proleev, Abe Ittycheriah, Soheil Hassas Yeganeh, George Polovets, Aleksandra Faust, Jiao Sun, Alban Rustemi, Pen Li, Rakesh Shivanna, Jeremiah Liu, Chris Welty, Federico Lebron, Anirudh Baddepudi, Sebastian Krause, Emilio Parisotto, Radu Soricut, Zheng Xu, Dawn Bloxwich, Melvin Johnson, Behnam Neyshabur, Justin Mao-Jones, Renshen Wang, Vinay Ramasesh, Zaheer Abbas, Arthur Guez, Constant Segal, Duc Dung Nguyen, James Svensson, Le Hou, Sarah York, Kieran Milan, Sophie Bridgers, Wiktor Gworek, Marco Tagliasacchi, James Lee-Thorp, Michael Chang, Alexey Guseynov, Ale Jakse Hartman, Michael Kwong, Ruizhe Zhao, Sheleem Kashem, Elizabeth Cole, Antoine Miech, Richard Tanburn, Mary Phuong, Filip Pavetic, Sebastian Cevey, Ramona Comanescu, Richard Ives, Sherry Yang, Cosmo Du, Bo Li, Zizhao Zhang, Mariko Iinuma, Clara Huiyi Hu, Aurko Roy, Shaan Bijwadia, Zhenkai Zhu, Danilo Martins, Rachel Saputro, Anita Gergely, Steven Zheng, Dawei Jia, Ioannis Antonoglou, Adam Sadovsky, Shane Gu, Yingying Bi, Alek Andreev, Sina Samangooei, Mina Khan, Tomas Kocisky, Angelos Filos, Chintu Kumar, Colton Bishop, Adams Yu, Sarah Hodgkinson, Sid Mittal, Premal Shah, Alexandre Moufarek, Yong Cheng, Adam Bloniarz, Jaehoon Lee, Pedram Pejman, Paul Michel, Stephen Spencer, Vladimir Feinberg, Xuehan Xiong, Nikolay Savinov, Charlotte Smith, Siamak Shakeri, Dustin Tran, Mary Chesus, Bernd Bohnet, George Tucker, Tamara von Glehn, Carrie Muir, Yiran Mao, Hideto Kazawa, Ambrose Slone, Kedar Soparkar, Disha Shrivastava, James Cobon-Kerr, Michael Sharman, Jay Pavagadhi, Carlos Araya, Karolis Misiunas, Nimesh Ghelani, Michael Laskin, David Barker, Qiuqia Li, Anton Briukhov, Neil Houlsby, Mia Glaese, Balaji Lakshminarayanan, Nathan Schucher, Yunhao Tang, Eli Collins, Hyeontaek Lim, Fangxiaoyu Feng, Adria Recasens, Guangda Lai, Alberto Magni, Nicola De Cao, Aditya Siddhant, Zoe Ashwood, Jordi Orbay, Mostafa Dehghani, Jenny Brennan, Yifan He, Kelvin Xu, Yang Gao, Carl Saroufim, James Molloy, Xinyi Wu, Seb Arnold, Solomon Chang, Julian Schrittwieser, Elena Buchatskaya, Soroush Radpour, Martin Polacek, Skye Giordano, Ankur Bapna, Simon Tokumine, Vincent Hellendoorn, Thibault Sottiaux, Sarah Cogan, Aliaksei Severyn, Mohammad Saleh, Shantanu Thakoor, Laurent Shefey, Siyuan Qiao, Meenu Gaba, Shuo yin Chang, Craig Swanson, Biao Zhang, Benjamin Lee, Paul Kishan Rubenstein, Gan Song, Tom Kwiatkowski, Anna Koop, Ajay Kanan, David Kao, Parker Schuh, Axel Stjerngren, Golnaz Ghiasi, Gena Gibson, Luke Vilnis, Ye Yuan, Felipe Tiengo Ferreira, Aishwarya Kamath, Ted Klimenko, Ken Franko, Kefan Xiao, Indro Bhattacharya, Miteyan Patel, Rui Wang, Alex Morris, Robin Strudel, Vivek Sharma, Peter Choy, Sayed Hadi Hashemi, Jessica Landon, Mara Finkelstein, Priya Jhakra, Justin Frye, Megan Barnes, Matthew Mauger, Dennis Daun, Khuslen Baatarsukh, Matthew Tung,

Wael Farhan, Henryk Michalewski, Fabio Viola, Felix de Chaumont Quiry, Charline Le Lan, Tom Hudson, Qingze Wang, Felix Fischer, Ivy Zheng, Elspeth White, Anca Dragan, Jean baptiste Alayrac, Eric Ni, Alexander Pritzel, Adam Iwanicki, Michael Isard, Anna Bulanova, Lukas Zilka, Ethan Dyer, Devendra Sachan, Srivatsan Srinivasan, Hannah Muckenhirn, Honglong Cai, Amol Mandhane, Mukarram Tariq, Jack W. Rae, Gary Wang, Kareem Ayoub, Nicholas FitzGerald, Yao Zhao, Woohyun Han, Chris Alberti, Dan Garrette, Kashyap Krishnakumar, Mai Gimenez, Anselm Levskaya, Daniel Sohn, Josip Matak, Inaki Iturrate, Michael B. Chang, Jackie Xi-ang, Yuan Cao, Nishant Ranka, Geoff Brown, Adrian Hutter, Vahab Mirrokni, Nanxin Chen, Kaisheng Yao, Zoltan Egyed, Francois Galilee, Tyler Liechty, Praveen Kallakuri, Evan Palmer, Sanjay Ghemawat, Jasmine Liu, David Tao, Chloe Thornton, Tim Green, Mimi Jasarevic, Sharon Lin, Victor Cotruta, Yi-Xuan Tan, Noah Fiedel, Hongkun Yu, Ed Chi, Alexander Neitz, Jens Heitkaemper, Anu Sinha, Denny Zhou, Yi Sun, Charbel Kaed, Brice Hulse, Swaroop Mishra, Maria Georgaki, Sneha Kudugunta, Clement Farabet, Izhak Shafran, Daniel Vlasic, Anton Tsitsulin, Rajagopal Ananthanarayanan, Alen Carin, Guolong Su, Pei Sun, Shashank V, Gabriel Carvajal, Josef Broder, Iulia Comsa, Alena Repina, William Wong, Warren Weilun Chen, Peter Hawkins, Egor Filonov, Lucia Loher, Christoph Hirsenschall, Weiyi Wang, Jingchen Ye, Andrea Burns, Hardie Cate, Diana Gage Wright, Federico Piccinini, Lei Zhang, Chu-Cheng Lin, Ionel Gog, Yana Kulizhskaya, Ashwin Sreevatsa, Shuang Song, Luis C. Cobo, Anand Iyer, Chetan Tekur, Guillermo Garrido, ZhuYun Xiao, Rupert Kemp, Huaixiu Steven Zheng, Hui Li, Ananth Agarwal, Christel Ngani, Kati Goshvadi, Rebeca Santamaria-Fernandez, Wojciech Fica, Xinyun Chen, Chris Gorgolewski, Sean Sun, Roopal Garg, Xinyu Ye, S. M. Ali Eslami, Nan Hua, Jon Simon, Pratik Joshi, Yelin Kim, Ian Tenney, Sahitya Potluri, Lam Nguyen Thiet, Quan Yuan, Florian Luisier, Alexandra Chronopoulou, Salvatore Scellato, Praveen Srinivasan, Minmin Chen, Vinod Koverkathu, Valentin Dalibard, Yaming Xu, Brennan Saeta, Keith Anderson, Thibault Sellam, Nick Fernando, Fantine Huot, Junehyuk Jung, Mani Varadarajan, Michael Quinn, Amit Raul, Maigo Le, Ruslan Habalov, Jon Clark, Komal Jalan, Kalesha Bullard, Achintya Singhal, Thang Luong, Boyu Wang, Sujeewan Rajayogam, Julian Eisenschlos, Johnson Jia, Daniel Finchelstein, Alex Yakubovich, Daniel Balle, Michael Fink, Sameer Agarwal, Jing Li, Dj Dvijotham, Shalini Pal, Kai Kang, Jaclyn Konzelmann, Jennifer Beattie, Olivier Dousse, Diane Wu, Remi Crocker, Chen Elkind, Siddhartha Reddy Jonnalagadda, Jong Lee, Dan Holtmann-Rice, Krystal Kallarackal, Rosanne Liu, Denis Vnukov, Neera Vats, Luca Invernizzi, Mohsen Jafari, Huanjie Zhou, Lilly Taylor, Jennifer Prendki, Marcus Wu, Tom Eccles, Tianqi Liu, Kavya Kopparapu, Francoise Beaufays, Christof Angermueller, Andreea Marzoca, Shourya Sarcar, Hilal Dib, Jeff Stanway, Frank Perbet, Nejc Trdin, Rachel Sterneck, Andrey Khorlin, Dinghua Li, Xihui Wu, Sonam Goenka, David

Madras, Sasha Goldshtein, Willi Gierke, Tong Zhou, Yaxin Liu, Yannie Liang, Anais White, Yunjie Li, Shreya Singh, Sanaz Bahargam, Mark Epstein, Sujoy Basu, Li Lao, Adnan Ozturk, Carl Crous, Alex Zhai, Han Lu, Zora Tung, Neeraj Gaur, Alanna Walton, Lucas Dixon, Ming Zhang, Amir Globerson, Grant Uy, Andrew Bolt, Olivia Wiles, Milad Nasr, Ilia Shumailov, Marco Selvi, Francesco Piccinno, Ricardo Aguilar, Sara McCarthy, Misha Khalman, Mrinal Shukla, Vlado Galic, John Carpenter, Kevin Villela, Haibin Zhang, Harry Richardson, James Martens, Matko Bosnjak, Shreyas Ram-mohan Belle, Jeff Seibert, Mahmoud Alnahlawi, Brian McWilliams, Sankalp Singh, Annie Louis, Wen Ding, Dan Popovici, Lenin Simicich, Laura Knight, Pulkit Mehta, Nishesh Gupta, Chongyang Shi, Saaber Fatehi, Jovana Mitrovic, Alex Grills, Joseph Pagadora, Tsendsuren Munkhdalai, Dessie Petrova, Danielle Eisenbud, Zhishuai Zhang, Damion Yates, Bhavishya Mittal, Niles Tripuraneni, Yanis Assael, Thomas Brovelli, Prateek Jain, Miha-jlo Velimirovic, Canfer Akbulut, Jiaqi Mu, Wolf-gang Macherey, Ravin Kumar, Jun Xu, Haroon Qureshi, Gheorghe Comanici, Jeremy Wiesner, Zhi-tao Gong, Anton Ruddock, Matthias Bauer, Nick Felt, Anirudh GP, Anurag Arnab, Dustin Zelle, Jonas Rothfuss, Bill Rosgen, Ashish Shenoy, Bryan Seybold, Xinjian Li, Jayaram Mudigonda, Goker Erdogan, Jiawei Xia, Jiri Simsa, Andrea Michi, Yi Yao, Christopher Yew, Steven Kan, Isaac Caswell, Carey Radebaugh, Andre Elisseeff, Pedro Valenzuela, Kay McKinney, Kim Paterson, Albert Cui, Eri Latorre-Chimoto, Solomon Kim, William Zeng, Ken Durden, Priya Ponnappalli, Tiberiu Sosea, Christopher A. Choquette-Choo, James Manyika, Brona Robenek, Harsha Vashisht, Sebastien Pereira, Hoi Lam, Marko Velic, Denese Owusu-Afriyie, Katherine Lee, Tolga Bolukbasi, Alicia Parrish, Shawn Lu, Jane Park, Balaji Venkatraman, Alice Talbert, Lambert Rosique, Yuchung Cheng, Andrei Sozanschi, Adam Paszke, Praveen Kumar, Jessica Austin, Lu Li, Khalid Salama, Bartek Perz, Wooyeol Kim, Nandita Dukkhipati, Anthony Baryshnikov, Christos Kapla-nis, XiangHai Sheng, Yuri Chervonyi, Caglar Unlu, Diego de Las Casas, Harry Askham, Kathryn Tun-yasuvunakool, Felix Gimeno, Siim Poder, Chester Kwak, Matt Miecnikowski, Vahab Mirrokni, Alek Dimitriev, Aaron Parisi, Dangyi Liu, Tomy Tsai, Toby Shevlane, Christina Kouridi, Drew Garmon, Adrian Goedeckemeyer, Adam R. Brown, Anitha Vi-jayakumar, Ali Elqursh, Sadegh Jazayeri, Jin Huang, Sara Mc Carthy, Jay Hoover, Lucy Kim, Sandeep Kumar, Wei Chen, Courtney Biles, Garrett Bingham, Evan Rosen, Lisa Wang, Qijun Tan, David Engel, Francesco Pongetti, Dario de Cesare, Dongseong Hwang, Lily Yu, Jennifer Pullman, Srini Narayanan, Kyle Levin, Siddharth Gopal, Megan Li, Asaf Aha-roni, Trieu Trinh, Jessica Lo, Norman Casagrande, Roopali Vij, Loic Matthey, Bramandia Ramadhana, Austin Matthews, CJ Carey, Matthew Johnson, Kre-mena Goranova, Rohin Shah, Shereen Ashraf, King-shuk Dasgupta, Rasmus Larsen, Yicheng Wang, Man-ish Reddy Vuyyuru, Chong Jiang, Joana Ijazi, Kazuki Osawa, Celine Smith, Ramya Sree Boppana, Tay-

- Ian Bilal, Yuma Koizumi, Ying Xu, Yasemin Altun, Nir Shabat, Ben Bariach, Alex Korchemniy, Kiam Choo, Olaf Ronneberger, Chimezie Iwuanyanwu, Shubin Zhao, David Soergel, Cho-Jui Hsieh, Irene Cai, Shariq Iqbal, Martin Sundermeyer, Zhe Chen, Elie Bursztein, Chaitanya Malaviya, Fadi Biadisy, Prakash Shroff, Inderjit Dhillon, Tejasi Latkar, Chris Dyer, Hannah Forbes, Massimo Nicosia, Vitaly Nikolaev, Somer Greene, Marin Georgiev, Pidong Wang, Nina Martin, Hanie Sedghi, John Zhang, Praseem Banzal, Doug Fritz, Vikram Rao, Xuezhi Wang, Jiageng Zhang, Viorica Patraucean, Dayou Du, Igor Mordatch, Ivan Jurin, Lewis Liu, Ayush Dubey, Abhi Mohan, Janek Nowakowski, Vlad-Doru Ion, Nan Wei, Reiko Tojo, Maria Abi Raad, Drew A. Hudson, Vaishakh Keshava, Shubham Agrawal, Kevin Ramirez, Zhichun Wu, Hoang Nguyen, Ji Liu, Madhavi Sewak, Bryce Pettrini, DongHyun Choi, Ivan Philips, Ziyue Wang, Ioana Bica, Ankush Garg, Jarek Wilkiewicz, Priyanka Agrawal, Xiaowei Li, Danhao Guo, Emily Xue, Naseer Shaik, Andrew Leach, Sadh MNM Khan, Julia Wiesinger, Sammy Jerome, Abhishek Chakladar, Alek Wenjiao Wang, Tina Ornduff, Folake Abu, Alireza Ghaffarkhah, Marcus Wainwright, Mario Cortes, Frederick Liu, Joshua Maynez, Andreas Terzis, Pouya Samangouei, Riham Mansour, Tomasz Kępa, François-Xavier Aubet, Anton Algymr, Dan Banica, Agoston Weisz, Andras Orban, Alexandre Senges, Ewa Andrejczuk, Mark Geller, Niccolo Dal Santo, Valentin Anklin, Majd Al Merey, Martin Baeuml, Trevor Strohmman, Junwen Bai, Slav Petrov, Yonghui Wu, Demis Hassabis, Koray Kavukcuoglu, Jeff Dean, and Oriol Vinyals. 2024. [Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context](#). *Preprint*, arXiv:2403.05530.
- Menno van Zaanen and Simon Zwarts. 2006. [Unsupervised measurement of translation quality using multi-engine, bi-directional translation](#). In *Proceedings of the 19th Australian Joint Conference on Artificial Intelligence: Advances in Artificial Intelligence*, AI'06, page 1208–1214, Berlin, Heidelberg. Springer-Verlag.
- 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 *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 6000–6010, Red Hook, NY, USA. Curran Associates Inc.
- Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. 2022. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*.
- Yiming Wang, Pei Zhang, Baosong Yang, Derek F. Wong, and Rui Wang. 2024. [Latent space chain-of-embedding enables output-free llm self-evaluation](#). *Preprint*, arXiv:2410.13640.
- Warren Weaver. 1952. Translation. In *Proceedings of the Conference on Mechanical Translation*.
- Haoran Xu, Young Jin Kim, Amr Sharaf, and Hany Hassan Awadalla. 2023. A paradigm shift in machine translation: Boosting translation performance of large language models. *arXiv preprint arXiv:2309.11674*.
- Ruijie Xu, Zengzhi Wang, Run-Ze Fan, and Pengfei Liu. 2024. [Benchmarking benchmark leakage in large language models](#). *Preprint*, arXiv:2404.18824.
- Xin Zhou, Martin Weyssow, Ratnadira Widayarsi, Ting Zhang, Junda He, Yunbo Lyu, Jianming Chang, Beiqi Zhang, Dan Huang, and David Lo. 2025. [Lessleak-bench: A first investigation of data leakage in llms across 83 software engineering benchmarks](#). *Preprint*, arXiv:2502.06215.

A Potential Risks

The potential risks associated with this work include the misuse of the generated artifacts outside of research contexts. To mitigate this, we have included a license in the appendix that specifies the intended use of the artifacts and restricts their use to research purposes only. We release them under the MIT license: <https://github.com/ulab-uiuc/consistencychecker>.

B Artifact Use

License compliance We confirm that our use of existing artifacts complies with their respective licenses. For artifacts created in this work, we explicitly specify intended use in the documentation, ensuring compatibility with original access conditions. Derivatives of data accessed for research purposes are restricted to research contexts only.

Model licenses Claude-3.5-Sonnet, GPT-4, GPT-4o-mini, and Gemini-1.5-Pro are all proprietary LLMs. In the Qwen family, Qwen-2.5-1.5B, Qwen-2.5-7B, Qwen-2.5-14B, and Qwen-2.5-32B are licensed under Apache-2.0, while Qwen-2.5-72B uses the Qwen license. Phi-3-Medium is under the MIT license. Mistral-Large is under the Mistral Research License. LLaMA-3-70B is licensed under the llama3 license, while LLaMA-3.1-8B and LLaMA-3.1-79B are under the llama3.1 license. nvidia/NV-Embed-v2, the embedding model we used, is under the license CC-BY-NC 4.0.

API compliance Open-source LLMs greater than 8B are accessed through Fireworks AI API. The Claude-3.5-Sonnet is accessed via official Anthropic endpoints. Gemini-1.5-Pro is visited from Gemini official API endpoints, and so is Mistral-Large, which is from Mistral official API. Phi-3-Medium is accessed through the NVIDIA Nim API. We accept and are compliant to each of their terms, respectively.

Benchmark use We do not use any existing benchmarks directly; we only reference their published results, and also the language pairs used in WMT 2024 automatic benchmarks.

C Documentation of Artifacts

We provide comprehensive documentation for the artifacts used and created in this work. This includes details on the domains, languages, and linguistic phenomena covered, as well as the demo-

graphic groups represented. This study involves the generation of benchmarks and evaluation log files involving English, German, French, Spanish, Czech, Ukrainian, Japanese, and Chinese.

D Statistics for Data

We report relevant statistics for the data used and created in this work. The four evaluator models involved in this study each generated 2 benchmark files, with one for the translation task and one for the AI-assisted programming task. Each of these benchmark files contains 10 self-consistent root nodes. For each LLM evaluated on each of these two tasks, it generated a self-consistency tree for each root node, and the degree and height of which are described in the paper.

E Model Size and Budget

Model size and configuration We report the number of parameters for each model used in this study:

Claude-3.5-Sonnet: unknown size;
GPT-4: unknown size;
GPT-4o-mini: unknown size;
Gemini-1.5-Pro: unknown size;
Mistral-Large: 128 billion parameters;
LLaMA-3-70B: 70 billion parameters;
Phi-3-Medium: 14 billion parameters;
Qwen-2.5-1.5B: 1.5 billion parameters;
Qwen-2.5-7B: 7 billion parameters;
Qwen-2.5-14B: 14 billion parameters;
Qwen-2.5-32B: 32.5 billion parameters;
Qwen-2.5-72B: 72 billion parameters;
LLaMA-3.1-8B: 8 billion parameters;
LLaMA-3.1-70B: 70 billion parameters.

Computational resources This project was conducted on 2 NVIDIA RTX A6000 GPUs. All LLMs with fewer than 8 billion parameters were hosted locally using vLLM on these GPUs, with random seed set to 42. Larger models and proprietary APIs (such as GPT-4o-mini, Qwen-2.5-14B, Qwen-2.5-70B, and LLaMA-3.1-72B) were accessed through external APIs.

Model variants All LLMs in the LLaMA and Qwen families used in this work are instruction-tuned versions. No additional quantization was applied to any model.

F Experimental Settings

We discuss the experimental setup, including hyperparameter search and the best-found hyperparameter values. The temperature for text generation was set to 0.6. The tree structure used for evaluation, on the two tasks, aside from the experiment for comparison with WMT 2024, had built 10 self-consistency tree with an out-degree of 3 (operation pairs) and a height of 3. These parameters were chosen based on preliminary experiments to balance diversity and consistency in the generated outputs. In the experiments to compare our framework with WMT 2024, for language pairs with English as one of the languages, the out-degree is 1 and the height of the trees is 12. For all those pairs without English as one of the two languages, the out-degree is 3 and the height is 3.

G Parameters for Packages

We used the following packages and reported their implementation, model, and parameter settings: `nlk.translate.bleu_score`: For calculating BLEU scores, we used the `sentence_bleu` function with `SmoothingFunction().method1`.

H Information about Use of AI Assistants

We acknowledge the use of AI assistants in the preparation of this work. The AI assistants were used for generating and refining text, as well as for providing suggestions on experimental design and hyperparameter tuning. All outputs generated by AI assistants were reviewed and validated by the authors to ensure accuracy and relevance.

I Additional Experimental Results

Additional results on path length We also release our results on these 2 defined tasks in these tables: Table 6, 7, and 8, where we report the numbers for $C_1(\mathcal{F})$, $C_2(\mathcal{F})$, and $C_3(\mathcal{F})$. They correspond to and provide more information than the tables Table 4 and 5. They also manifest the trend that under the same circumstances $C_2(\mathcal{F})$ is lower than $C_1(\mathcal{F})$, and $C_3(\mathcal{F})$ is lower than $C_2(\mathcal{F})$.

Additional results on correlation scores in WMT 2024 Table 9 shows the comparison between our ConsistencyChecker rankings and WMT 2024 rankings. In most cases, at least 4 out of 6 LLMs share the same order across both evaluations. To further quantify the relationship, we calculate Pearson’s correlation between our consistency scores

and WMT’s three automated metrics. Since AutoRank and MetricX are inversely related to performance, we use their absolute values for the correlation analysis. As illustrated in Figure 5, there is a strong correlation between ConsistencyChecker scores and WMT 2024 evaluations. This relationship is evident using both NV-Embed-v2-based cosine embedding similarity and BLEU as the similarity metric.

J Additional Case Study

J.1 Full Examples of Leaf Nodes

We include the full example of one root node and two leaf nodes generated by LLaMA-3.1-8B and LLaMA-3.1-70B for consistency evaluation in Figure 6.

J.2 Full Examples of Self-consistency Tree

In this section, we present two complete self-consistency trees with a depth of 3 and a branching factor of 3 for the machine translation task. These trees are generated by two different models: Qwen-2.5-7B and Qwen-2.5-72B.

This is the full content of the self-consistency tree generated by Qwen-2.5-7B on machine translation tasks.

Root (Level 0, NV-Embed-v2 similarity to root 1.0000). In the ever-evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker

of the shifting tides in international finance. The ECB’s decision to recalibrate its quantitative easing (QE) program, alongside a series of nuanced interest rate adjustments, has sent shockwaves through the financial markets, particularly affecting emerging economies that are intricately linked to the

Eurozone’s economic health. These changes, while designed to stabilize the Euro

and stimulate economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and heightened risk aversion among investors. The ECB’s forward guidance, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address the diverse and dynamic economic conditions across different regions. In response, smaller financial institutions and central banks in emerging markets have been compelled to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to insulate their economies from external shocks and capitalize on the emerging opportunities. The interplay between these global and local economic strategies underscores the complex and

Table 6: **Machine translation consistency evaluation with different path lengths.** Scores(%) evaluate LLMs’ ability to attain semantics after translating across different languages.

Evaluatee	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$
GPT-4o-mini	99.2 ± 0.0	98.7 ± 0.0	98.0 ± 0.0
Qwen-2.5-1.5B	91.9 ± 0.1	86.5 ± 0.1	80.3 ± 0.5
Qwen-2.5-7B	96.4 ± 0.1	93.5 ± 0.3	90.0 ± 0.8
Qwen-2.5-14B	97.4 ± 0.0	96.1 ± 0.1	94.7 ± 0.1
Qwen-2.5-32B	97.6 ± 0.1	97.4 ± 0.0	96.4 ± 0.0
Qwen-2.5-72B	99.1 ± 0.0	98.3 ± 0.0	97.2 ± 0.0
LLaMA-3.1-8B	89.5 ± 0.1	75.7 ± 1.6	67.5 ± 3.0
LLaMA-3.1-70B	87.4 ± 0.7	80.7 ± 1.1	71.9 ± 3.2

Table 7: **AI-assisted programming consistency evaluation with different path lengths.** Scores(%) evaluate LLMs’ ability to preserve functionality after multi-step algorithm transformation.

Evaluatee	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$
GPT-4o-mini	90.6 ± 0.2	84.7 ± 0.6	76.5 ± 2.7
Qwen-2.5-1.5B	90.2 ± 0.2	80.0 ± 0.6	63.4 ± 1.4
Qwen-2.5-7B	85.4 ± 0.2	80.2 ± 0.2	71.7 ± 0.4
Qwen-2.5-14B	90.3 ± 0.2	83.1 ± 0.6	79.9 ± 1.0
Qwen-2.5-32B	91.1 ± 0.4	88.3 ± 0.7	85.1 ± 1.1
Qwen-2.5-72B	90.8 ± 0.2	85.3 ± 0.5	77.0 ± 1.9
LLaMA-3.1-8B	88.1 ± 0.2	76.0 ± 0.5	60.4 ± 1.0
LLaMA-3.1-70B	87.9 ± 0.4	82.7 ± 0.6	83.5 ± 1.0

Table 8: **Consistency results under different evaluator models and different path lengths.** (Top) Machine translation consistency scores (%) across model scales (7B–72B) and transform complexity ($C_1(\mathcal{F})$ – $C_3(\mathcal{F})$). (Bottom) Code equivalence preservation scores (%) for algorithm transformations at varying abstraction levels ($C_1(\mathcal{F})$ – $C_3(\mathcal{F})$).

Evaluator	Qwen-2.5-7B			Qwen-2.5-72B		
	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$
Machine translation						
Qwen-2.5-7B	90.1 ± 0.9	82.8 ± 3.0	76.8 ± 5.1	96.4 ± 0.1	93.5 ± 0.3	90.0 ± 0.8
Qwen-2.5-72B	90.0 ± 1.8	84.4 ± 4.2	78.2 ± 7.3	99.1 ± 0.0	98.3 ± 0.0	97.2 ± 0.0
LLaMA-3.1-8B	81.6 ± 0.5	73.0 ± 1.3	61.7 ± 4.8	89.5 ± 0.1	75.7 ± 1.6	67.5 ± 3.0
LLaMA-3.1-70B	82.3 ± 1.8	68.8 ± 2.3	30.8 ± 9.2	87.4 ± 0.7	80.7 ± 1.1	71.9 ± 3.2
AI-assisted programming						
Qwen-2.5-7B	86.2 ± 0.4	78.2 ± 0.8	72.2 ± 1.9	85.4 ± 0.2	80.2 ± 0.2	71.7 ± 0.4
Qwen-2.5-72B	87.2 ± 0.2	80.1 ± 0.6	79.4 ± 2.3	90.8 ± 0.2	85.3 ± 0.5	77.0 ± 1.9
LLaMA-3.1-8B	84.4 ± 0.2	72.1 ± 0.4	61.0 ± 1.1	88.1 ± 0.2	76.0 ± 0.5	60.4 ± 1.0
LLaMA-3.1-70B	91.0 ± 0.5	85.6 ± 1.0	81.5 ± 2.1	87.9 ± 0.4	82.7 ± 0.6	83.5 ± 1.0

interconnected nature of the modern financial system, where the decisions of one major player can have profound implications for the rest of the world.

Node 0 (Level 1, NV-Embed-v2 similarity to root 0.3336). The communication and forward strategies between global and local economic strategies highlight the complexity and interconnectivity of modern financial systems, where one player’s decisions can have far-reaching impacts on the rest of the world.

Node 0-0 (Level 2, NV-Embed-v2 similarity to root 0.2802). The communication and promotion strategy between global and local economic strategies highlights the complexity and interconnection of modern financial systems, where the decisions of a single actor can have reverberating impacts on the rest of the world.

Node 0-0-0 (Level 3, NV-Embed-v2 similarity to root 0.2820). The communication and promotion strategies between global and local economic strategies highlight the complexity and interconnectedness of modern financial systems, where the decisions of a single actor can have repercussions on the rest of the world.

Node 0-0-1 (Level 3, NV-Embed-v2 similarity to root 0.2400). The communications and propaganda strategy between global and local economic

strategies illustrates the complexity and the internal connection of modern financial systems, where the decisions of a single actor can affect the rest of the world.

Node 0-0-2 (Level 3, NV-Embed-v2 similarity to root 0.2693). The communication and promotion strategy between global and local economic strategies highlights the complexity and interconnection of modern financial systems, where the decisions of a single actor can have resonant effects on the rest of the world.

Node 0-1 (Level 2, NV-Embed-v2 similarity to root 0.2774). The communication and foresight strategies between global and local business strategies illustrate the complexity and interdependence of modern financial systems, where the decisions of one player can have far-reaching effects on the rest of the world.

Node 0-1-0 (Level 3, NV-Embed-v2 similarity to root 0.2717). The communication and foresight strategy between global and local business strategies illustrates the complexity and interdependence of modern financial systems, where the decisions of one player can have repercussions on the rest of the world.

Node 0-1-1 (Level 3, NV-Embed-v2 similarity to root 0.2615). The communication and prediction strategies between global and regional

Table 9: **Comprehensive results on machine translation tasks.** Consistency scores (%) across model and transform complexity ($C_1(\mathcal{F})$ - $C_3(\mathcal{F})$) using embedding similarity and BLEU scores, versus metrics in WMT 2024 benchmarks. The rows are arranged using WMT rankings.

Czech-Ukrainian	Embedding			BLEU			AutoRank↓	MetricX↓	CometKiwi↑
	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$			
Claude-3.5-Sonnet	99.4	98.9	98.1	87.1	80.4	67.3	1.7	1.0	0.683
GPT-4	98.7	97.8	96.4	78.1	69.2	55.5	2.0	1.4	0.677
Gemini-1.5-Pro	99.2	98.5	97.5	86.0	78.3	67.6	2.0	1.2	0.668
Mistral-Large	98.6	97.6	96.5	78.3	68.8	56.7	2.3	1.6	0.666
LLaMA-3-70B	96.4	96.7	95.9	79.2	70.4	55.8	2.6	2.0	0.661
Phi-3-Medium	71.1	58.3	44.9	28.9	14.6	5.6	9.1	6.5	0.425
English-German	Embedding			BLEU			AutoRank↓	MetricX↓	CometKiwi↑
	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$			
GPT-4	99.4	99.2	99.1	85.8	83.5	82.4	1.8	1.4	0.700
Claude-3.5-Sonnet	99.8	99.8	99.8	95.6	94.8	94.4	1.9	1.4	0.695
Mistral-Large	99.4	99.3	99.2	88.5	86.9	85.5	2.0	1.5	0.694
Gemini-1.5-Pro	99.8	99.7	99.6	93.3	92.8	91.4	2.2	1.5	0.688
LLaMA-3-70B	99.3	99.1	98.9	88.4	86.2	84.3	2.5	1.7	0.686
Phi-3-Medium	98.5	97.9	97.2	81.2	76.2	72.7	3.4	2.0	0.657
English-Japanese	Embedding			BLEU			AutoRank↓	MetricX↓	CometKiwi↑
	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$			
Claude-3.5-Sonnet	99.4	99.2	99.0	89.2	87.3	85.8	1.5	2.3	0.744
Gemini-1.5-Pro	99.3	99.2	99.1	88.1	86.3	84.9	1.7	2.5	0.734
GPT-4	98.4	98.1	97.8	76.6	72.2	69.4	1.7	2.7	0.740
LLaMA-3-70B	93.5	90.5	87.9	71.4	65.4	60.0	2.6	3.5	0.714
Phi-3-Medium	97.4	96.1	95.3	65.5	57.1	51.8	2.8	3.6	0.709
Mistral-Large	98.2	97.4	96.9	79.2	73.9	70.7	2.9	3.8	0.707
English-Chinese	Embedding			BLEU			AutoRank↓	MetricX↓	CometKiwi↑
	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$			
Claude-3.5-Sonnet	99.5	99.4	99.3	89.8	88.7	86.9	1.7	3.0	0.703
Gemini-1.5-Pro	99.4	99.3	99.3	89.4	88.5	86.9	1.8	3.1	0.698
GPT-4	98.6	98.3	98.1	76.4	73.3	71.4	2.0	3.3	0.693
Mistral-Large	97.5	97.7	97.5	82.5	79.5	77.6	2.8	4.0	0.665
LLaMA-3-70B	98.6	98.3	97.9	82.4	78.8	75.9	2.8	3.9	0.662
Phi-3-Medium	98.1	97.3	96.6	72.2	65.9	61.8	3.1	4.2	0.648
Japanese-Chinese	Embedding			BLEU			AutoRank↓	MetricX↓	CometKiwi↑
	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$	$C_1(\mathcal{F})$	$C_2(\mathcal{F})$	$C_3(\mathcal{F})$			
Claude-3.5-Sonnet	98.7	97.7	96.0	77.4	65.4	49.6	1.7	3.5	0.603
Gemini-1.5-Pro	98.8	98.1	96.8	80.1	70.0	54.0	1.9	3.5	0.595
GPT-4	97.6	96.0	93.7	65.3	51.6	37.3	2.1	3.8	0.597
LLaMA-3-70B	96.2	93.4	88.9	64.6	48.0	28.5	3.1	4.7	0.578
Mistral-Large	98.1	96.7	94.4	69.4	55.6	39.9	3.5	4.9	0.568
Phi-3-Medium	93.9	89.3	84.9	52.1	38.6	26.0	4.0	5.1	0.552

business concepts illustrate the complexity and interconnectedness of modern financial systems, where the decisions of one player can have far-reaching effects on the rest of the world.

Node 0-1-2 (Level 3, NV-Embed-v2 similarity to root 0.2721). The communication and forecast strategy between global and local business strategies illustrate the complexity and interdependence of modern financial systems, where the decisions of one player can have long-lasting effects on the rest of the world.

Node 0-2 (Level 2, NV-Embed-v2 similarity to root 0.2783). The strategy of communication and advancement between global and local strategies highlights the complexity and interconnection of modern financial systems, where the decisions of a single player can have long-term repercussions for the rest of the world.

Node 0-2-0 (Level 3, NV-Embed-v2 similarity to root 0.2595). The communication and advancement strategy between global and local strategies highlights the complexity and interconnection of

modern financial systems, where the decisions of a single player can have long-term repercussions for the rest of the world.

Node 0-2-1 (Level 3, NV-Embed-v2 similarity to root 0.2852). The strategy of communication and progress balancing between global and local strategies illustrates the complexity and connection of modern financial systems, where the decisions of a single player can have long-term effects on the rest of the world.

Node 0-2-2 (Level 3, NV-Embed-v2 similarity to root 0.2687). The strategy of communication and advancement between global and local strategies highlights the complexity and interconnection of modern financial systems, in which the decisions of a single player can have long-term repercussions for the rest of the world.

Node 1 (Level 1, NV-Embed-v2 similarity to root 0.9537). In the constantly changing landscape of global economic policy, the recent measures of the European Central Bank (ECB) stand out as a significant sign of evolving

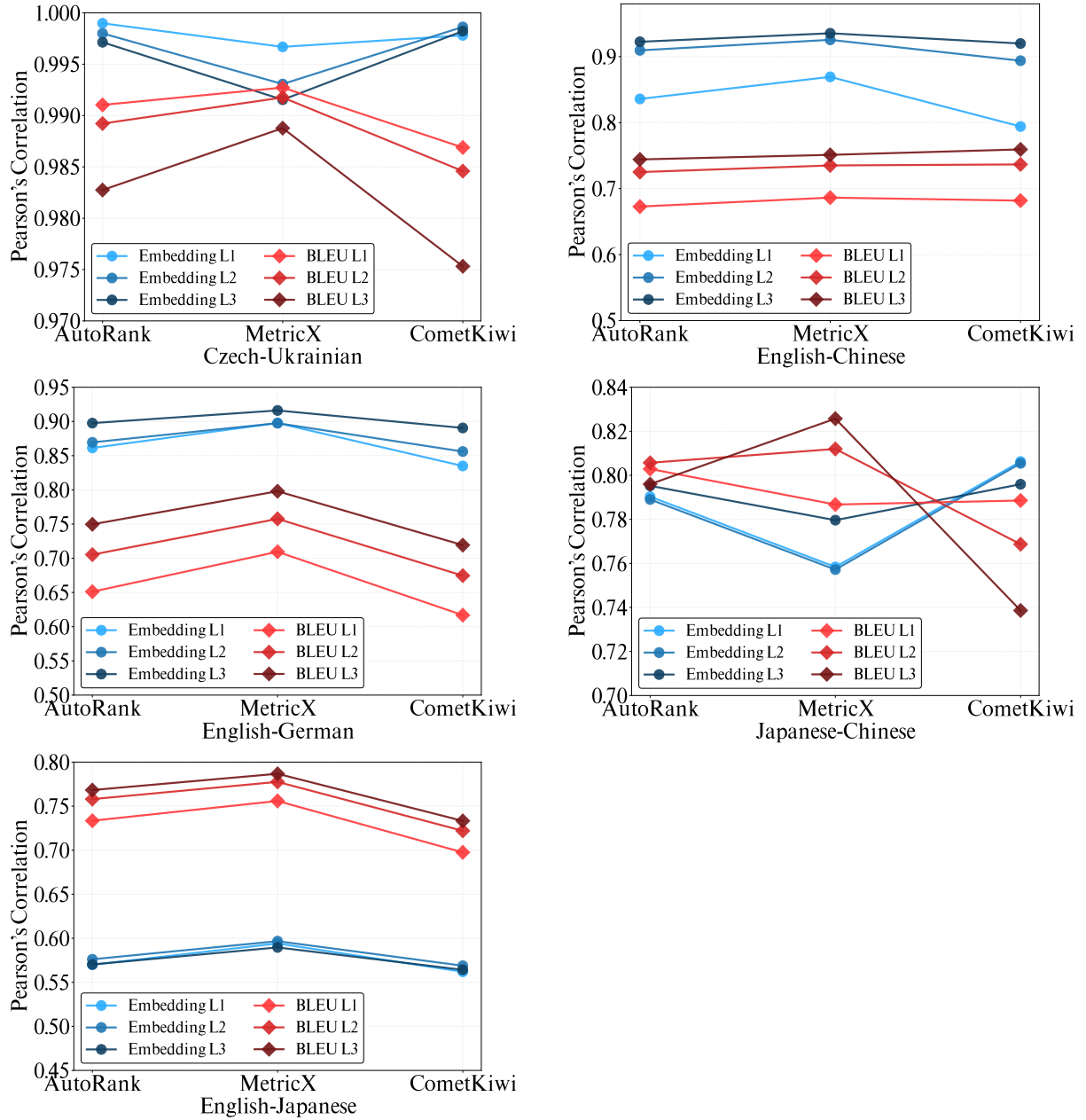


Figure 5: **Correlation between ConsistencyChecker evaluation and metrics in WMT 2024 benchmarks.** The three figures on the left column from top to bottom correspond to the Czech-Ukrainian, English-German, and English-Japanese language pairs. The two figures on the right column from top to bottom correspond to the English-Chinese and Japanese-Chinese language pairs. Each of these figures shows the correlation between the ConsistencyChecker scores by embedding and by BLEU at $C_1(\mathcal{F})$, $C_2(\mathcal{F})$, and $C_3(\mathcal{F})$, with the three metrics used in WMT 2024, AutoRank, MetricX, and CometKiwi. The results show that both embedding-based and BLEU-based ConsistencyChecker scores share high correlation with the three metrics in WMT 2024.

directions in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) programs and a series of nuanced interest rate adjustments have sparked shockwaves in financial markets, particularly in emerging economies closely linked to the economic health of the eurozone. These changes, aimed at stabilizing the euro and fostering growth in the European Union, have far-reaching consequences, including capital flight, currency volatility, and heightened risk aversion among investors.

The forward guidance of the ECB, which seeks to provide clarity and predictability, is met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to accommodate the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in emerging economies are forced to take a range of adaptive measures, from tightening their interest conditions to diversifying their investment portfolios, to shield their economies

Root Node: def main(): return "In the rapidly evolving landscape of artificial intelligence (AI), the integration of these advanced technologies into healthcare systems represents a significant frontier with profound implications for medical diagnostics and patient care. Recent advancements have demonstrated the potential of AI to enhance diagnostic accuracy, streamline clinical workflows, and personalize treatment plans, thereby improving patient outcomes and reducing healthcare costs. However, the deployment of AI in this sensitive domain is not without its challenges. Ethical considerations, such as data privacy, algorithmic bias, and the potential for job displacement, are paramount. Moreover, the regulatory landscape must adapt to ensure that AI applications meet stringent safety and efficacy standards. Despite these hurdles, the collaboration between tech developers, healthcare providers, and regulatory bodies is fostering a cautious yet optimistic approach to AI integration. The future of healthcare may well be defined by the seamless fusion of human expertise and machine intelligence, creating a hybrid model that leverages the strengths of both to deliver more effective and efficient care."

LLaMA-3.1-8B Leaf Node: def main(): return "In the constantly evolving scenario of artificial intelligence (AI), the integration of these advanced technologies in medical systems represents a significant frontier with important implications for medical diagnosis and patient care. Recent advances have demonstrated the potential of AI to improve diagnostic accuracy, simplify clinical workflows, and personalize treatment plans, leading to better patient outcomes and a reduction in medical costs. However, implementing AI in this sensitive domain does not come without its challenges. Ethical considerations, such as data privacy, algorithmic bias, and job displacement, are fundamental. Furthermore, the regulatory landscape must adapt to ensure that AI applications meet strict standards of safety and efficacy. Despite these difficulties, collaboration between technology developers, healthcare providers, and regulatory bodies is promoting a cautious yet optimistic approach to the integration of AI. The future of healthcare may be defined by the seamless fusion of human experience and artificial intelligence, creating a hybrid model that leverages the strengths of both to offer more effective and efficient care."

LLaMA-3.1-70B Leaf Node: def main(): return "In the rapidly evolving landscape of artificial intelligence (AI), the integration of these advanced technologies into healthcare systems represents a significant frontier with profound implications for medical diagnosis and patient care. Recent advances have demonstrated AI's potential to improve diagnostic accuracy, streamline clinical workflows, and personalize treatment plans, thereby enhancing patient outcomes and reducing healthcare costs. However, the deployment of AI in this sensitive domain is not without challenges. Ethical considerations, such as data privacy, algorithmic bias, and potential job displacement, are paramount. Moreover, the regulatory landscape must adapt to ensure that AI applications meet stringent safety and efficacy standards. Despite these obstacles, collaboration between technology developers, healthcare providers, and regulatory agencies is fostering a cautious yet optimistic approach to AI integration. The future of healthcare may be defined by the seamless fusion of human expertise and machine intelligence, creating a hybrid model that leverages the strengths of both to deliver more effective and efficient care."

Figure 6: Example of root nodes and leaf nodes in a self-consistency tree.

from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of modern financial systems, where the decisions of a major player can have profound impacts on the rest of

the world.

Node 1-0 (Level 2, NV-Embed-v2 similarity to root 0.9234). In the constantly evolving global economic landscape, the recent measures of the European Central Bank (ECB) stand out as a significant sign of emerging directions in the

international financial world. The ECB's decision to modulate its Quantitative Easing (QE) programs and a series of subtle adjustments to interest rates have had repercussions in financial markets, particularly in emerging economies closely linked to the economic health of the eurozone. These changes, aimed at stabilizing the euro and promoting growth in the European Union, have consequences far beyond, including capital flight, currency volatility, and an increased risk aversion among investors. The ECB's forward guidance, which aims to provide clarity and predictability, is met with mixed reactions, as some market analysts suggest that it lacks the flexibility needed to adapt to diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging economies are forced to adopt a range of adaptive measures, from reducing interest rates to diversifying investment portfolios, in order to protect their economies from external shocks and seize emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of modern financial systems, where the decisions of a major player can have profound impacts on the rest of the world.

Node 1-0-0 (Level 3, NV-Embed-v2 similarity to root 0.0520). None

Node 1-0-1 (Level 3, NV-Embed-v2 similarity to root 0.8577). In the constantly changing global economic landscape, the recent measures of the European Central Bank (ECB) represent an important signal for the emerging directions on the international financial market. The decision of the ECB to modulate its Quantitative Liquidity Provision (QLD) and make a series of subtle adjustments to interest rates has elicited reactions from the financial market, particularly in emerging economies that rely on the economic condition of the Eurozone. These changes, aimed at stabilizing the Euro and promoting economic growth in the European Union, have consequences far beyond the Eurozone region, including capital flight, currency fluctuations, and increased risk aversion among investors. The ECB's foresight in providing clarity and predictability is met with mixed reactions, as some market analysts believe it lacks the flexibility to adapt to diverse and dynamic economic conditions in various regions. In response to this, smaller financial institutions and central banks in emerging economies must take a series of adequate measures, ranging from reducing interest rates to diversifying investment portfolios, to protect their economies from external shocks and to capitalize on emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of modern financial systems, in which the decisions of a major player can profoundly influence the world economy.

Node 1-0-2 (Level 3, NV-Embed-v2 similarity to root 0.8955). In the constant scenario of global economic change, the recent measures of the European Central Bank (ECB) stand out as a significant sign of emerging directions in international finance. The ECB's decision to modulate its Quantitative Easing (QE) programs

and a series of subtle adjustments to interest rates have had repercussions on financial markets, especially in emerging countries closely linked to the economic health of the eurozone. These changes, aimed at stabilizing the euro and promoting growth in the European Union, have consequences beyond, including capital flight, currency volatility, and increased risk aversion among investors. The guidance provided by the ECB for the future, with the aim of providing clarity and predictability, is received with mixed reactions, as some market analysts suggest that it lacks the flexibility needed to adapt to dynamic and diverse economic conditions in different regions. As a response, smaller financial institutions and central banks of emerging countries are obliged to adopt a variety of adaptive measures, from reducing interest rates to diversifying investment portfolios, in order to protect their economies from external shocks and seize emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of modern financial systems, where the decisions of a major player can have profound impacts on the rest of the world.

Node 1-1 (Level 2, NV-Embed-v2 similarity to root 0.9165). In the constantly changing landscape of global economic and political conditions, the recent measures of the European Central Bank (ECB) stand out as a significant indicator of developing trends in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) programs and implement a series of fine-tuned interest rate adjustments triggers shockwaves in financial markets, especially in emerging economies closely linked to the economic health of the Eurozone. These changes, aimed at stabilizing the euro and promoting growth in the European Union, have far-reaching consequences, including capital flight, currency optimality, and increased risk aversion among investors. The forward-looking but cautious approach of the ECB, which provides clarity and predictability, meets with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to accommodate the diverse and dynamic economic conditions in various regions. In response, smaller financial institutions and central banks in emerging economies are forced to take a range of adaptive measures, from interest rate reticence to diversifying their investment portfolios, in order to shield their economies from external shocks and seize new opportunities. The interaction between these global and local economic strategies highlights the complex and intertwined nature of modern financial systems, where the decisions of one major player can have profound impacts on the rest of the world.

Node 1-1-0 (Level 3, NV-Embed-v2 similarity to root 0.9070). In the constantly evolving global economic and political landscape, the recent measures of the European Central Bank (ECB) stand out as a significant indicator of trends developing in the international financial world. The decision of the ECB to modify its Quantitative Easing (QE) programs and apply a series of delicate interest rate adjustments has repercussions in financial markets, particularly in emerging economies closely linked to the

economic health of the eurozone. These changes, aimed at stabilizing the euro and fostering growth in the European Union, have far-reaching consequences, including capital flight, currency optimization, and increased investor aversion. The ECB's forward-looking and prudent approach, which brings clarity and predictability, meets with mixed reactions, as some market analysts claim it lacks sufficient flexibility to adapt to the diverse and dynamic economic conditions of various regions. In response, smaller financial institutions and central banks of emerging economies are compelled to adopt a range of adaptive measures, from interest rate restraint to portfolio diversification, to protect their economies from external shocks and seize new opportunities. The interaction between these global and local economic strategies highlights the complex and interwoven nature of modern financial systems, where the decisions of a major player can have profound impacts on the rest of the world.

Node 1-1-1 (Level 3, NV-Embed-v2 similarity to root 0.8573). In the constantly changing landscape of global economic and political conditions, the recent measures of the European Central Bank (ECB) serve as an important indicator of developing trends in the international financial world. The decision of the ECB to adjust its Quantitative Easing programs and implement a series of nuanced interest rate adjustments has waves in the financial markets, especially in development-assisted economies closely linked to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and promoting growth in the European Union, have far-reaching consequences, including capital flight, currency optimality, and increased risk aversion among investors. The ECB's advancing but cautious approach, which offers clarity and predictability, garners mixed reactions, as some market analysts argue that it lacks the necessary flexibility to take into account the diverse and dynamic economic conditions in various regions. As a result, smaller financial institutions and central banks in development-assisted economies must take a series of adaptive measures, ranging from interest rate restraint to diversification of their investment portfolios, to protect their economies from external shocks and seize new opportunities. The interaction between these global and local economic strategies highlights the complex and interwoven nature of modern financial systems, where the decisions of a key player can have strong impacts on the rest of the world.

Node 1-1-2 (Level 3, NV-Embed-v2 similarity to root 0.9194). In the ever-changing landscape of the global economic and political scenario, the recent measures of the European Central Bank (ECB) stand out as a significant indicator of emerging trends in the international financial world. The decision of the ECB to adjust its Quantitative Easing (QE) programs and implement a series of refined interest rate adjustments triggers shockwaves in financial markets, particularly in emerging economies closely linked to the economic health of the Eurozone. These changes, aimed at stabilizing the euro and fostering growth in the European Union, have far-reaching consequences,

including capital flight, currency optimization, and increased risk aversion among investors. The forward-looking but cautious approach of the ECB, which provides clarity and predictability, receives mixed reactions, as some market analysts argue that it lacks the necessary flexibility to adapt to dynamically and diversely changing economic conditions in various regions. As a response, smaller financial institutions and central banks of emerging economies are compelled to adopt a series of adaptive measures, ranging from reluctance to adjust interest rates to diversifying their investment portfolios, in order to protect their economies from external shocks and seize new opportunities. The interaction between these global and local economic strategies highlights the complex and intertwined nature of modern financial systems, where the decisions of one key player can have a profound impact on the rest of the world.

Node 1-2 (Level 2, NV-Embed-v2 similarity to root 0.9204). In the constantly evolving global landscape of economic policy, the recent measures of the European Central Bank (ECB) stand out as a significant sign of evolving directions in the international financial world. The decision of the ECB to adjust its quantitative easing (QE) programs and a series of subtle interest rate adjustments have sparked shocks in financial markets, especially in emerging economies closely tied to the economic health of the eurozone. These changes, aimed at stabilizing the euro and promoting growth in the European Union, have far-reaching consequences, including capital flight, currency volatility, and increased risk aversion among investors. The future orientation of the ECB, which seeks to provide clarity and predictability, is met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address the diverse and dynamic economic conditions in different regions. As a response, small financial institutions and central banks in emerging economies are compelled to take a range of adaptive measures, from tightening interest rate conditions to diversifying their investment portfolios, to protect their economies from external shocks and seize emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of modern financial systems, where decisions by a major player can have profound impacts on the rest of the world.

Node 1-2-0 (Level 3, NV-Embed-v2 similarity to root 0.8980). In the constantly evolving global economic landscape, the recent measures of the European Central Bank (ECB) stand out as a significant sign of emerging directions in the international financial world. The ECB's decision to modify its quantitative easing (QE) programs and a series of subtle adjustments to interest rates have caused ripples in financial markets, particularly in emerging economies closely linked to the economic health of the eurozone. These changes, aimed at stabilizing the euro and promoting development within the European Union, have far-reaching consequences, including capital flight, currency volatility, and increased risk aversion among investors. The future orientation of the ECB, which seeks to bring clarity and

predictability, is met with mixed reactions, as some market analysts believe it lacks the necessary flexibility to respond to diverse and dynamic economic conditions in different regions. In response, financial institutions and central banks in emerging economies are forced to take a range of adaptive measures, from tightening interest rate conditions to diversifying their investment portfolios, to protect their economies against external shocks and seize emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of modern financial systems, where the decisions of a major actor can have deep impacts on the rest of the world.

Node 1-2-1 (Level 3, NV-Embed-v2 similarity to root 0.9313). In the constantly changing global landscape of international finance and economic policy, the recent measures of the European Central Bank (ECB) serve as an important indicator of evolving developments. The ECB's decision to adjust its Quantitative Easing (QE) programs and implement a series of subtle interest rate adjustments has sparked shocks in financial markets, particularly in emerging economies that have close ties to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and fostering economic growth in the European Union, have far-reaching consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The forward-looking orientation of the ECB, which seeks to provide clear and predictable measures, is met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address the diverse and dynamic economic conditions in different regions. In response, small financial institutions and central banks in emerging economies are obligated to take a range of adaptive measures, from revealing interest rates to diversifying their investment portfolios, to protect their economies from external shocks and seize new opportunities. The interaction between these global and local economic strategies highlights the complex and intertwined nature of modern financial systems, where the decisions of a major player can have profound impacts on the rest of the world.

Node 1-2-2 (Level 3, NV-Embed-v2 similarity to root 0.9081). In the constantly evolving global economic landscape, recent measures by the European Central Bank (ECB) stand out as a significant sign of changing directions in the international financial world. The ECB's decision to adjust its quantitative easing (QE) programs and a series of subtle adjustments in interest rates have caused shocks in financial markets, especially in emerging countries closely linked to the economic health of the eurozone. These changes, aimed at stabilizing the euro and promoting growth in the European Union, have far-reaching consequences, including capital flight, currency volatility, and increased risk aversion among investors. The ECB's future orientation, seeking to provide clarity and predictability, is received with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address diverse and dynamic economic conditions in different regions. As a response, small financial institutions and central banks of

emerging countries are compelled to take a series of adaptive measures, from tightening interest rate conditions to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of modern financial systems, where the decisions of one major player can have a profound impact on the rest of the world.

Node 2 (Level 1, NV-Embed-v2 similarity to root 0.9469). In the constantly evolving landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in international financial waters. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging countries closely linked to the economic health of the Eurozone. These changes, designed to stabilize the euro and stimulate economic growth within the European Union, have had broader implications, including capital flight, currency volatility, and an increase in risk aversion among investors. The ECB's forward guidance, intended to provide clarity and predictability, has received mixed reactions, as some market analysts argue that it lacks sufficient flexibility to address the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in emerging countries have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in order to insulate themselves from external shocks and capitalize on emerging opportunities. The interplay between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where the decisions of one major player can have profound implications for the rest of the world.

Node 2-0 (Level 2, NV-Embed-v2 similarity to root 0.9227). In the constantly evolving global economic landscape, the recent policy adjustments of the European Central Bank (ECB) stand out as a significant marker of changes in international financial waters. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle adjustments to interest rates, has caused ripples in financial markets, particularly affecting emerging countries closely linked to the economic health of the eurozone. These changes, designed to stabilize the euro and stimulate economic growth within the European Union, have broader implications, including capital flight, currency volatility, and increased risk aversion among investors. The forward guidance, aimed at providing clarity and predictability, has received mixed reactions, as some market analysts believe it lacks the necessary flexibility to address the complex and changing economic conditions in different regions. In response, smaller financial institutions and central banks in emerging countries have had to take a series of adaptive measures, from tightening monetary

policy to diversifying portfolios, to protect themselves from external shocks and seize emerging opportunities. The interaction between global and regional economic strategies highlights the complexity and interconnectivity of modern financial systems, where the decisions of one major player can have far-reaching effects on the rest of the world.

Node 2-0-0 (Level 3, NV-Embed-v2 similarity to root 0.9002). In the constantly evolving global economic landscape, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant sign of changes on the horizon for international financial waters. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has caused ripples in financial markets, particularly affecting emerging economies closely linked to the economic health of the eurozone. These changes, designed to stabilize the euro and stimulate economic growth within the European Union, have broader implications, including capital flight, currency volatility, and increased investor aversion. The future direction, aimed at bringing clarity and predictability, has received mixed reactions, as some market analysts believe it lacks the necessary flexibility to respond to complex and changing economic conditions in different regions. In response, small financial institutions and central banks of emerging economies have had to take a series of adaptive measures, ranging from tightening monetary policy to diversifying portfolios, to protect themselves from external shocks and seize emerging opportunities. The interaction between global and regional economic strategies highlights the complexity and interconnectedness of modern financial systems, where the decisions of a major actor can have reverberating effects on the rest of the world.

Node 2-0-1 (Level 3, NV-Embed-v2 similarity to root 0.9184). In the constantly changing global economic landscape, the recent policy changes of the European Central Bank (ECB) are a significant marker for changes in international financial waters. The ECB's decision to recalibrate its Quantitative Easing (QE) program, along with a series of subtle adjustments to interest rates, has sent waves through the financial markets, particularly in Emerging Markets that are closely linked to the economic well-being of the Eurozone. These changes, aimed at stabilizing the Euro and promoting economic development within the European Union, have broader implications, including capital flight, currency volatility, and increased risk aversion among investors. The forward guidance strategy, which seeks to provide clarity and predictability, is met with mixed reactions as some market analysts believe it lacks the necessary flexibility to be fair to complex and changing economic conditions in various regions. In response, smaller financial institutions and central banks in Emerging Markets have taken a range of adaptive measures, from abstaining from monetary policy to diversifying portfolios, to protect themselves from external shocks and seize new opportunities. The interaction between global and regional economic strategies highlights the complexity and interdependence of modern financial

systems, where the decisions of a major player can have far-reaching effects on the rest of the world.

Node 2-0-2 (Level 3, NV-Embed-v2 similarity to root 0.8992). In the constant scenario of the evolving global economic landscape, the recent political adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in international financial waters. The decision of the ECB to recalibrate its quantitative easing (QE) program, along with a series of subtle adjustments in interest rates, has caused ripples in financial markets, particularly affecting emerging countries that are closely linked to the economic health of the eurozone. These changes, designed to stabilize the euro and stimulate economic growth within the European Union, have broader implications, including capital flight, exchange rate volatility, and increased risk aversion among investors. The forward-looking orientation, aimed at providing clarity and predictability, has received mixed reactions, as some market analysts believe it lacks the necessary flexibility to address changing and complex economic conditions in different regions. As a response, smaller financial institutions and central banks of emerging countries have had to take a series of adaptive measures, ranging from tightening monetary policy to diversifying portfolios, to protect themselves from external shocks and seize emerging opportunities. The interaction between global and regional economic strategies highlights the complexity and interconnectivity of modern financial systems, where the decisions of one major player can have far-reaching effects on the rest of the world.

Node 2-1 (Level 2, NV-Embed-v2 similarity to root 0.9394). In the constantly evolving landscape of global economic governance, the most recent political adjustment of the European Central Bank (ECB) stands as an important landmark for changes in the international financial waters. The decision of the ECB to recalibrate its Quantitative Easing (QE) program, as well as a series of slight interest rate adjustments, have unleashed shockwaves in the financial markets, particularly in emerging countries with close ties to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and strengthening economic growth within the European Union, have far-reaching implications, including capital flight, currency volatility, and an increase in risk aversion among investors. The ECB's forecasts, which are intended to provide clarity and predictability, have received mixed reactions, as some market analysts argue that they do not offer enough flexibility to account for the diverse and dynamic economic conditions in various regions. In response, smaller financial institutions and central banks in emerging countries have taken a range of adaptive measures, from lowering monetary policies to diversifying their investment portfolios, to insulate themselves from external shocks and take advantage of new opportunities. The interaction between these global and local economic strategies underscores the complex and intertwined nature of the modern financial system, where the decisions of one major player can have significant implications for the rest of the world.

Node 2-1-0 (Level 3, NV-Embed-v2 similarity to root 0.9184). In the constantly evolving global economic landscape, the recent adjustment of the monetary policy of the European Central Bank (ECB) serves as an important landmark for changes in international financial waters. The decision of the ECB to recalibrate its Quantitative Easing (QE) program, along with a series of minor adjustments to interest rates, has caused ripples in financial markets, particularly among emerging economies closely linked to the economic health of the eurozone. These changes, aimed at stabilizing the euro and strengthening economic growth in the European Union, have wide-ranging implications, including capital flight, currency volatility, and an increased risk appetite among investors. The ECB's forecasts, which aim to bring clarity and predictability, have received mixed reactions, as some market analysts believe they do not provide enough flexibility to account for diverse and dynamic economic conditions in different regions. In response, financial institutions and central banks of emerging economies have adopted a range of adaptive measures, from easing monetary policies to diversifying investment portfolios, to protect against external shocks and take advantage of new opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of the modern financial system, where the decisions of a major player can have significant implications for the rest of the world.

Node 2-1-1 (Level 3, NV-Embed-v2 similarity to root 0.8603). In the constantly changing landscape of global economic and political governance, the latest political change of the European Central Bank (ECB) represents an important milestone for changes in international financial waters. The decision of the ECB to recalibrate its Quantitative Easing (QE) program, as well as a series of minor interest rate adjustments, has triggered waves of shockwaves in the financial market, particularly in development countries that rely on the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and strengthening economic growth in the European Union, have far-reaching implications, including capital flight, currency volatility, and an increase in risk aversion among investors. The ECB's forecasts, which aim to increase clarity and predictability, have received mixed reactions, as some market analysts argue that they are not sufficiently flexible to accommodate the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in developing countries have taken a series of adaptable measures, ranging from reducing monetary policy to diversifying their investment portfolios, to insulate themselves from external shocks and seize new opportunities. The interaction between these global and local economic strategies underscores the complex and interwoven nature of the modern financial system, where the decisions of one major player can have significant impacts on the rest of the world.

Node 2-1-2 (Level 3, NV-Embed-v2 similarity to root 0.9302). In the constant scenario of the evolution of global economic governance, the latest political adjustment by the European Central Bank (ECB) has risen to become an important

milestone for changes in international financial waters. The ECB's decision to recalibrate its Quantitative Easing (QE) program, along with a series of minor adjustments to interest rates, has triggered shockwaves in financial markets, especially in emerging economies with close ties to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and strengthening economic growth within the European Union, have long-term implications, including capital flight, currency volatility, and an increase in risk aversion among investors. The ECB's projections, which seek to provide clarity and predictability, have received mixed reactions, as some market analysts argue that they do not offer sufficient flexibility to address the diverse and dynamic economic conditions in various regions. In response, financial institutions and central banks of emerging economies have taken a series of adaptive measures, from lowering their monetary policies to diversifying their investment portfolios, to protect themselves from external shocks and seize new opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of the modern financial system, where decisions by a single dominant player can have significant implications for the rest of the world.

Node 2-2 (Level 2, NV-Embed-v2 similarity to root 0.9398). In the evolving landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in international financial waters. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle adjustments in interest rates, has sent shockwaves through financial markets, particularly affecting emerging countries closely linked to the economic health of the European Economic Area. These changes, designed to stabilize the euro and stimulate economic growth within the European Union, have broader implications, including capital flight, currency volatility, and an increase in risk aversion among investors. The future guidance provided by the ECB, intended to offer clarity and predictability, has received mixed reactions, as some market analysts argue that it lacks sufficient flexibility to address the dynamic and diverse economic conditions in different regions. In response, smaller financial institutions and central banks of emerging countries have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, with the aim of insulating themselves from external shocks and capitalizing on emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of the modern financial system, where the decisions of a single major player can have profound implications for the rest of the world.

Node 2-2-0 (Level 3, NV-Embed-v2 similarity to root 0.9372). In the constantly evolving landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in international financial

waters. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle adjustments to interest rates, sent shockwaves through financial markets, particularly affecting emerging economies closely tied to the economic health of the European Economic Area. These changes, designed to stabilize the euro and stimulate economic growth in the European Union, have broader implications, including currency volatility and an increase in risk aversion among investors. The future guidance provided by the ECB, aimed at offering clarity and predictability, has received mixed reactions, as some financial analysts believe it lacks flexibility to respond to dynamic and diverse economic conditions in different regions. In response, smaller financial institutions and central banks of emerging economies have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, with the goal of insulating these countries from external shocks and capitalizing on emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of the modern financial system, where the decisions of a single major actor can have profound implications for the rest of the world.

Node 2-2-1 (Level 3, NV-Embed-v2 similarity to root 0.9168). In the constantly evolving scenario of global economic and political governance, the recent political adjustment by the European Central Bank (ECB) represents a significant feature of changes in the international financial waters. The ECB's decision to recalibrate its Quantitative Easing (QE) program, as well as a series of subtle adjustments to interest rates, have triggered shockwaves in the financial markets, particularly in emerging countries closely linked to the economic health of the European Economic and Monetary Union (EMU). These changes, aimed at stabilizing the euro and promoting economic development in the European Union, have broader impacts, including capital flight, currency instability, and increased risk aversion among investors. The future guidance provided to the ECB to offer clarity and predictability has elicited mixed reactions, as some market analysts argue that it does not offer sufficient flexibility to account for the dynamic and diverse economic conditions in various regions. In response, smaller financial institutions and central banks in emerging countries have taken a range of adaptive measures, from the abstinence of monetary policy to the diversification of their investor portfolios, with the goal of insulating themselves from external shocks and capitalizing on emerging opportunities. The interaction between these global and local economic strategies highlights the complex and interconnected nature of the modern financial system, where decisions of a single major player can have far-reaching impacts on the rest of the world.

Node 2-2-2 (Level 3, NV-Embed-v2 similarity to root 0.9388). In the evolutionary landscape of global economic governance, the recent political adjustments made by the European Central Bank

(ECB) stand out as a significant marker of changes in international financial waters. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle adjustments in interest rates, has sent shockwaves through financial markets, particularly affecting emerging countries that are closely linked to the economic health of the Euro Area. These changes, designed to stabilize the euro and stimulate economic growth within the European Union, have broader implications, including capital flight, currency volatility, and increased risk aversion among investors. The guidance provided by the ECB, intended to offer clarity and predictability, has received mixed reactions, as some market analysts argue that it lacks sufficient flexibility to address the dynamic and diverse economic conditions in different regions. In response, smaller financial institutions and central banks of emerging countries have been forced to adopt a series of adaptive measures, ranging from tightening monetary policies to diversifying their investment portfolios, with the goal of insulating themselves from external shocks and capitalizing on emerging opportunities. The interaction between these global and local economic strategies illustrates the complex and interconnected nature of the modern financial system, where the decisions of a single major player can have profound implications for the rest of the world.

This is the full content of the self-consistency tree generated by Qwen-2.5-72B on machine translation tasks.

Root (Level 0, NV-Embed-v2 similarity to root 1.0000). In the ever-evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of the shifting tides in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, alongside a series of nuanced interest rate adjustments, has sent shockwaves through the financial markets, particularly affecting emerging economies that are intricately linked to the Eurozone's economic health. These changes, while designed to stabilize the Euro and stimulate economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and heightened risk aversion among investors. The ECB's forward guidance, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address the diverse and dynamic economic conditions across different regions. In response, smaller financial institutions and central banks in emerging markets

have been compelled to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to insulate their economies from external shocks and capitalize on the emerging opportunities. The interplay between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where the decisions of one major player can have profound implications for the rest of the world.

Node 0 (Level 1, NV-Embed-v2 similarity to root 0.9745). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, coupled with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and stimulate economic growth within the European Union, have had long-term consequences, including capital flight, currency volatility, and increased risk aversion among investors. The ECB's forward guidance, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions across different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 0-0 (Level 2, NV-Embed-v2 similarity to root 0.9672). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, coupled with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly impacting emerging economies closely tied to the economic health of the eurozone. Although these changes are designed to stabilize the euro and boost economic growth within the European Union, they have had long-term consequences, including capital flight, currency volatility, and increased risk aversion among investors. Future guidance from the ECB, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions across different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to adopt a range of adaptive

measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 0-0-0 (Level 3, NV-Embed-v2 similarity to root 0.9672). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, coupled with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies closely tied to the economic health of the eurozone. Although these changes are designed to stabilize the euro and stimulate economic growth within the European Union, they have had long-term consequences, including capital flight, currency volatility, and increased risk aversion among investors. The ECB's future guidance, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions across different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to adopt a range of adaptation measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 0-0-1 (Level 3, NV-Embed-v2 similarity to root 0.9424). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a significant indicator of trends in the international financial world. The ECB's decision to recalibrate its Quantitative Easing (QE) program, coupled with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly to emerging economies that are closely tied to the economic health of the Eurozone. While these changes are designed to stabilize the Euro and promote economic growth within the European Union, they have had long-term consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The future guideline of the ECB, which aims to create clarity and predictability, has been met with mixed reactions, with some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to adopt a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to

protect their economies from external shocks and capitalize on new opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a key player can have far-reaching impacts on the rest of the world.

Node 0-0-2 (Level 3, NV-Embed-v2 similarity to root 0.9677). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly impacting emerging economies closely tied to the economic health of the eurozone. While these changes are designed to stabilize the euro and stimulate economic growth within the European Union, they have had long-term consequences, including capital flight, exchange rate volatility, and an increase in risk aversion among investors. The ECB's future guidance, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, financial institutions and smaller central banks in emerging markets have been forced to adopt a series of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 0-1 (Level 2, NV-Embed-v2 similarity to root 0.9274). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a significant indicator of trends in the international financial world. The ECB's decision to recalibrate its Quantitative Easing (QE) program, coupled with a series of subtle interest rate changes, has sent shockwaves through financial markets, particularly to emerging countries closely tied to the economic health of the Eurozone. While these changes aim to stabilize the Euro and promote economic growth within the European Union, they have long-term consequences, including capital flight, currency volatility, and increased risk aversion among investors. The future leadership of the ECB, which is intended to provide clarity and predictability, has been met with mixed reactions, with some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on opportunities. The interaction between these

global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching effects on the rest of the world.

Node 0-1-0 (Level 3, NV-Embed-v2 similarity to root 0.9488). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) are a significant indicator of trends in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of small interest rate changes, has sent shockwaves through financial markets, particularly to emerging countries closely tied to the economic health of the eurozone. While these changes aim to stabilize the euro and promote economic growth within the European Union, they have long-term consequences, including capital outflows, exchange rate fluctuations, and increased risk aversion among investors. The future direction of the ECB, which is supposed to provide clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have long-term effects on the rest of the world.

Node 0-1-1 (Level 3, NV-Embed-v2 similarity to root 0.9263). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a significant indicator of trends in the international financial world. The ECB's decision to recalibrate its Quantitative Easing (QE) program, coupled with a series of subtle interest rate changes, has sent shockwaves through financial markets, particularly in emerging countries that are closely tied to the economic health of the Eurozone. While these changes aim to stabilize the euro and promote economic growth within the European Union, they have long-term consequences, including capital flight, currency volatility, and increased risk aversion among investors. The future leadership of the ECB, which is expected to provide clarity and predictability, has been met with mixed reactions, with some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to implement a range of adjustment measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player

can have far-reaching impacts on the rest of the world.

Node 0-1-2 (Level 3, NV-Embed-v2 similarity to root 0.9382). In the constantly changing landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) are a significant indicator of trends in international financial markets. The ECB's decision to recalibrate its Quantitative Easing (QE) program, along with a series of subtle changes in interest rates, has sent shockwaves through financial markets, particularly to emerging countries closely tied to the economic health of the Eurozone. While these changes aim to stabilize the Euro and promote economic growth within the European Union, they have long-term consequences, including capital flight, currency volatility, and an increase in risk aversion among investors. The future leadership of the ECB, which is intended to provide clarity and predictability, has been met with mixed reactions, with some market analysts arguing that it lacks the necessary flexibility to respond to diverse and dynamic economic conditions in different regions. In response, financial institutions and smaller central banks in emerging markets have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching effects on the rest of the world.

Node 0-2 (Level 2, NV-Embed-v2 similarity to root 0.9748). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and stimulate economic growth within the European Union, have had long-term consequences, including capital flight, currency volatility, and an increase in risk aversion among investors. The ECB's forward guidance, intended to provide clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, financial institutions and smaller central banks in emerging markets have been forced to adopt a variety of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications

for the rest of the world.

Node 0-2-0 (Level 3, NV-Embed-v2 similarity to root 0.9653). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, coupled with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and boost economic growth within the European Union, have had long-term consequences, including capital outflows, exchange rate volatility, and increased risk aversion among investors. The ECB's future guidance, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, financial institutions and central banks of emerging markets have been forced to adopt a variety of adaptive measures, ranging from tightening monetary policies to diversifying their investment portfolios, in an effort to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 0-2-1 (Level 3, NV-Embed-v2 similarity to root 0.9697). In the constantly changing landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant indicator of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly in emerging countries that are closely tied to the economic health of the Eurozone. These changes, although designed to stabilize the Euro and promote economic growth within the European Union, had long-term consequences, including capital flight, currency fluctuations, and an increase in investor risk aversion. The ECB's forward guidance, intended to provide clarity and predictability, has been met with mixed reactions. Some market analysts argue that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, financial institutions and smaller central banks in emerging markets have been forced to adopt a range of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on new opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching effects on the rest of the world.

Node 0-2-2 (Level 3, NV-Embed-v2 similarity

to root 0.9670). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of trends in international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent ripples through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and stimulate economic growth within the European Union, have had long-term consequences, including capital flight, currency volatility, and an increase in risk aversion among investors. The future guidance of the ECB, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to diverse and dynamic economic conditions in different regions. In response, financial institutions and smaller central banks in emerging markets have been forced to adopt a variety of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in an effort to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 1 (Level 1, NV-Embed-v2 similarity to root 0.9453). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a significant sign of the shifting currents in international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differentiated interest rate adjustments, has sent shockwaves through the financial markets, particularly affecting emerging economies that are closely tied to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and boosting economic growth in the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The forward-looking communication of the ECB, intended to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to accommodate the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to shield their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching impacts on the rest of the world.

Node 1-0 (Level 2, NV-Embed-v2 similarity to

root 0.9745). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) are a significant sign of changing tides in international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differentiated interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, aimed at stabilizing the euro and stimulating economic growth within the European Union, have had significant consequences, including capital outflows, exchange rate fluctuations, and increased risk aversion among investors. The forward guidance from the ECB, intended to provide clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to adapt to diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have significant impacts on the rest of the world.

Node 1-0-0 (Level 3, NV-Embed-v2 similarity to root 0.9586). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) are a significant sign of changes in the currents of international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differentiated interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies closely tied to the economic health of the eurozone. These changes, aimed at stabilizing the euro and stimulating economic growth within the European Union, have had significant consequences, including capital outflows, exchange rate fluctuations, and increased risk aversion among investors. The future guidance of the ECB, intended to bring clarity and predictability, has been met with mixed reactions, some market analysts arguing that they lack the necessary flexibility to adapt to the diverse and dynamic economic conditions in different regions. In response, small financial institutions and central banks of developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have significant impacts on the rest of the world.

Node 1-0-1 (Level 3, NV-Embed-v2 similarity to root 0.9468). In the constantly changing

environment of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a significant sign of changes in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differentiated interest rate adjustments, has sent shockwaves through the financial markets, especially the emerging economies that are closely tied to the economic health of the Eurozone. These changes, aimed at stabilizing the euro and boosting economic growth within the European Union, have had substantial consequences, including capital outflows, exchange rate fluctuations, and increased risk aversion among investors. The forward-looking guidance from the ECB, intended to provide clarity and predictability, has been met with mixed reactions, with some market analysts arguing that it lacks the necessary flexibility to adapt to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on existing opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a key player can have significant impacts on the rest of the world.

Node 1-0-2 (Level 3, NV-Embed-v2 similarity to root 0.9678). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) are a significant sign of changes in the tides of international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differentiated interest rate adjustments, has sent shock waves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, aimed at stabilizing the euro and stimulating economic growth within the European Union, have had significant consequences, including capital flows, exchange rate fluctuations, and an increase in risk aversion among investors. The ECB's forward guidance, intended to provide clarity and predictability, has been met with mixed reactions, with some market analysts arguing that it lacks the necessary flexibility to adapt to diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and take advantage of available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have significant impacts on the rest of the world.

Node 1-1 (Level 2, NV-Embed-v2 similarity to root 0.9418). In the constantly changing landscape of global economic governance, the recent political

adjustments by the European Central Bank (ECB) are a significant sign of the shifting tides in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of different interest rate adjustments, has sent shockwaves through the financial markets, particularly the emerging economies that are closely tied to the economic stability of the Eurozone. These changes, aimed at stabilizing the euro and promoting economic growth in the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The forward-looking communication of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to account for the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a key player can have far-reaching impacts on the rest of the world.

Node 1-1-0 (Level 3, NV-Embed-v2 similarity to root 0.9603). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) are a significant sign of trends changing in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of different interest rate adjustments, has sent shockwaves through financial markets, particularly emerging economies that are closely tied to the economic stability of the eurozone. These changes, aimed at stabilizing the euro and promoting economic growth within the European Union, have had significant consequences, including capital outflows, exchange rate fluctuations, and increased risk aversion among investors. The forward guidance from the ECB, which aims to provide clarity and predictability, has received mixed reactions, with some market analysts arguing that it lacks the necessary flexibility to account for the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and take advantage of available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a key player can have significant impacts on the rest of the world.

Node 1-1-1 (Level 3, NV-Embed-v2 similarity to root 0.9420). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank

(ECB) are a significant sign of the shifting currents in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of interest rate changes, has sent shockwaves through financial markets, particularly through emerging markets that are closely tied to the economic stability of the Eurozone. These changes, aimed at stabilizing the euro and promoting economic growth in the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The forward-looking communication of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to accommodate the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a key player can have far-reaching impacts on the rest of the world.

Node 1-1-2 (Level 3, NV-Embed-v2 similarity to root 0.9449). In the constantly changing landscape of global economic governance, recent political adjustments by the European Central Bank (ECB) are a significant sign of changes in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of different interest rate adjustments, has sent shockwaves through financial markets, particularly to emerging economies that are closely tied to the economic stability of the Eurozone. These changes, aimed at stabilizing the euro and promoting economic growth in the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and an increase in risk aversion among investors. The forward guidance of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to account for the diverse and dynamic economic conditions in different regions. In response, financial institutions and smaller central banks in developing countries have been forced to implement a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and take advantage of available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a key player can have widespread impacts on the rest of the world.

Node 1-2 (Level 2, NV-Embed-v2 similarity to root 0.9690). In the constantly changing landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) are a significant sign of the shifting

tides in international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differentiated interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and promoting economic growth in the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and an increase in risk aversion among investors. The forward-looking communication of the ECB, intended to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to adapt to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of developing countries have been forced to implement a variety of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching impacts on the rest of the world.

Node 1-2-0 (Level 3, NV-Embed-v2 similarity to root 0.9738). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) are a significant sign of changing tides in international finance. The ECB's decision to alter its Quantitative Easing (QE) program, coupled with a series of differentiated interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies closely tied to the economic health of the eurozone. These changes, aimed at stabilizing the euro and promoting economic growth in the European Union, have had significant consequences, including capital outflows, currency fluctuations, and increased risk aversion among investors. The ECB's forward guidance, intended to bring clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to adapt to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in developing countries have been forced to implement a variety of adaptive measures, ranging from monetary policy tightening to diversification of their investment portfolios, to protect their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have significant impacts on the rest of the world.

Node 1-2-1 (Level 3, NV-Embed-v2 similarity to root 0.9372). In the constantly changing structure of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a clear sign of the shifting

tides in the international financial world. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differentiated interest rate adjustments, has sent shockwaves through financial markets, particularly to emerging countries that are closely linked to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and promoting economic growth in the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The forward-looking communication of the ECB, intended to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to adapt to the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and the central banks of developing countries have been forced to implement a variety of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and to benefit from available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching impacts on the rest of the world.

Node 1-2-2 (Level 3, NV-Embed-v2 similarity to root 0.9695). In the constantly changing landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) are a significant sign of the shifting tides in international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of differential interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the Eurozone. These changes, aimed at stabilizing the Euro and promoting economic growth in the European Union, have had far-reaching consequences, including capital flight, monetary fluctuations, and an increase in risk aversion among investors. The forward-looking communication from the ECB, intended to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to adapt to diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of developing countries have been forced to implement a variety of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on available opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching impacts on the rest of the world.

Node 2 (Level 1, NV-Embed-v2 similarity to root 0.9723). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of changes

in the currents of international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are intricately linked to the economic health of the Eurozone. These changes, though designed to stabilize the Euro and stimulate economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and increased risk aversion among investors. The future orientation of the ECB, which seeks to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging markets have been forced to adopt a series of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in order to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 2-0 (Level 2, NV-Embed-v2 similarity to root 0.9709). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in the currents of international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and stimulate economic growth within the European Union, have had far-reaching consequences, including capital flight, exchange rate volatility, and increased risk aversion among investors. The future direction of the ECB, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, small financial institutions and central banks in emerging markets have been forced to adopt a series of adaptive measures, ranging from easing monetary policies to diversifying their investment portfolios, in order to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 2-0-0 (Level 3, NV-Embed-v2 similarity to root 0.9686). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in the currents of international finance. The ECB's

decision to recalibrate its quantitative easing (QE) program, coupled with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and boost economic growth within the European Union, have had broader consequences, including capital flight, exchange rate volatility, and increased risk aversion among investors. The future direction of the ECB, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, small financial institutions and central banks in emerging markets have been forced to adopt a series of adaptation measures, ranging from easing monetary policies to diversifying their investment portfolios, in order to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 2-0-1 (Level 3, NV-Embed-v2 similarity to root 0.9361). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a significant sign of changes in the currents of international financial markets. The ECB's decision to realign its Quantitative Easing (QE) program, coupled with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly in emerging countries that are closely tied to the economic health of the Eurozone. These changes, although designed to stabilize the Euro and promote economic growth within the European Union, have had far-reaching consequences, including capital flight, exchange rate volatility, and increased risk aversion among investors. The future direction of the ECB, which aims for clarity and predictability, has been met with mixed reactions. Some market analysts argue that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, small financial institutions and central banks in emerging markets have been forced to adopt a range of adaptive measures, from loosening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on new opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching implications for the rest of the world.

Node 2-0-2 (Level 3, NV-Embed-v2 similarity to root 0.9701). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in the currents of international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with

a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and stimulate economic growth within the European Union, have had far-reaching consequences, including capital flight, exchange rate volatility, and an increase in risk aversion among investors. The future direction of the ECB, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the various and dynamic economic conditions in different regions. In response, small financial institutions and central banks in emerging markets have been forced to adopt a series of adaptive measures, ranging from easing monetary policies to diversifying their investment portfolios, in order to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 2-1 (Level 2, NV-Embed-v2 similarity to root 0.9395). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) stand out as a significant sign of changes in the currents of international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through the financial markets, particularly the emerging economies that are closely tied to the economic health of the Eurozone. These changes, although designed to stabilize the euro and promote economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and increased risk aversion among investors. The future orientation of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to account for the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging countries have been forced to adopt a series of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching implications for the rest of the world.

Node 2-1-0 (Level 3, NV-Embed-v2 similarity to root 0.9659). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant sign of changes in the currents of international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves

through financial markets, particularly emerging economies closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and promote economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and increased risk aversion among investors. The future direction of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to account for diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging countries have been forced to adopt a series of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where the decisions of a major player can have far-reaching implications for the rest of the world.

Node 2-1-1 (Level 3, NV-Embed-v2 similarity to root 0.9380). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) are a significant sign of changes in the currents of international finance. The ECB's decision to adjust its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly through the emerging economies that are closely tied to the economic health of the Eurozone. These changes, although aimed at stabilizing the euro and promoting economic growth within the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The future orientation of the ECB, which is intended to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to consider the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging countries have been forced to adopt a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and seize new opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have far-reaching implications for the rest of the world.

Node 2-1-2 (Level 3, NV-Embed-v2 similarity to root 0.9696). In the constantly changing landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant sign of changes in the currents of international finance. The ECB's decision to adjust its Quantitative Easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets,

particularly to emerging economies that are closely linked to the economic health of the Eurozone. These changes, although designed to stabilize the euro and promote economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and an increase in risk aversion among investors. The future guidance of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to account for the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging countries have been forced to adopt a series of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and take advantage of emerging opportunities. The interaction between these global and local economic strategies underscores the complex and intertwined nature of the modern financial system, where decisions by a major player can have far-reaching implications for the rest of the world.

Node 2-2 (Level 2, NV-Embed-v2 similarity to root 0.9803). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in the currents of international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are intricately linked to the economic health of the Eurozone. These changes, although designed to stabilize the Euro and stimulate economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and an increase in risk aversion among investors. The future guidance of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging markets have been forced to adopt a series of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in order to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 2-2-0 (Level 3, NV-Embed-v2 similarity to root 0.9769). In the constantly evolving landscape of global economic governance, the recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in the currents of international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, combined with a series of subtle interest rate adjustments, has sent shock waves through financial markets,

particularly affecting emerging economies that are closely tied to the economic health of the eurozone. These changes, although designed to stabilize the euro and boost economic growth within the European Union, have had significant consequences, including capital flight, currency volatility, and increased risk aversion among investors. The ECB's future guidance, aimed at providing clarity and predictability, has been met with mixed reactions, some market analysts arguing that it lacks the necessary flexibility to respond to the diverse and dynamic economic conditions in different regions. In response, small financial institutions and central banks in emerging markets have been forced to adopt a series of adaptive measures, ranging from tightening monetary policies to diversifying their investment portfolios, in order to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.

Node 2-2-1 (Level 3, NV-Embed-v2 similarity to root 0.9269). In the constantly changing landscape of global economic governance, the recent political adjustments by the European Central Bank (ECB) stand out as a significant sign of changes in the flows of international finance. The ECB's decision to recalibrate its Quantitative Easing (QE) program, along with a series of subtle interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are closely tied to the economic health of the Eurozone. These changes, although designed to stabilize the euro and boost economic growth within the European Union, have had far-reaching consequences, including capital flight, currency fluctuations, and increased risk aversion among investors. The future leadership of the ECB, which is supposed to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to handle the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks of emerging markets have been forced to adopt a range of adaptive measures, from tightening monetary policy to diversifying their investment portfolios, to protect their economies from external shocks and capitalize on new opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a key player can have far-reaching implications for the rest of the world.

Node 2-2-2 (Level 3, NV-Embed-v2 similarity to root 0.9806). In the constantly evolving landscape of global economic governance, recent policy adjustments by the European Central Bank (ECB) stand out as a significant marker of changes in the currents of international finance. The ECB's decision to recalibrate its quantitative easing (QE) program, along with a series of subtle

interest rate adjustments, has sent shockwaves through financial markets, particularly affecting emerging economies that are intricately linked to the economic health of the Eurozone. These changes, although designed to stabilize the Euro and stimulate economic growth within the European Union, have had far-reaching consequences, including capital flight, currency volatility, and an increase in risk aversion among investors. The future guidance of the ECB, which aims to provide clarity and predictability, has been met with mixed reactions, as some market analysts argue that it lacks the necessary flexibility to address the diverse and dynamic economic conditions in different regions. In response, smaller financial institutions and central banks in emerging markets have been forced to adopt a series of adaptive measures, from tightening monetary policies to diversifying their investment portfolios, in order to protect their economies from external shocks and capitalize on emerging opportunities. The interaction between these global and local economic strategies underscores the complex and interconnected nature of the modern financial system, where decisions by a major player can have profound implications for the rest of the world.