

# No Significant Relationship Between Conceptual Depth and Graph Distance to Concrete Letter Nodes in the Copycat Slipnet

Slipnet Analysis Project  
slipnet\_analysis/

February 1, 2026

## Abstract

The Copycat system, developed by Douglas Hofstadter and Melanie Mitchell, employs a semantic network called the slipnet where each node has a “conceptual depth” parameter intended to capture its level of abstraction. We investigate whether conceptual depth correlates with the topological distance (hop count) from abstract concept nodes to concrete letter nodes (a–z). Using breadth-first search on an undirected graph representation of the slipnet, we computed minimum hop distances for 33 non-letter nodes, assigning unreachable nodes a penalty distance of  $2 \times \max(\text{hops})$ . Statistical analysis reveals no significant correlation between conceptual depth and hop count (Pearson  $r = 0.281$ ,  $p = 0.113$ ; Spearman  $\rho = 0.141$ ,  $p = 0.433$ ). The coefficient of determination ( $R^2 = 0.079$ ) indicates that conceptual depth explains only 7.9% of the variance in hop distance. These findings demonstrate that conceptual depth and network topology are orthogonal design dimensions in the Copycat architecture.

## 1 Introduction

The Copycat project, developed by Douglas Hofstadter and Melanie Mitchell in the 1980s and 1990s, represents a landmark effort in computational cognitive science to model analogical reasoning (1; 2). The system operates on letter-string analogy problems of the form “if abc changes to abd, what does ppqrr change to?” While the domain is deliberately simple, the un-

derlying cognitive architecture embodies sophisticated principles about how concepts are represented and manipulated during reasoning.

### 1.1 The Slipnet Architecture

Central to Copycat’s operation is the *slipnet*, a semantic network containing 59 nodes representing concepts relevant to the letter-string domain. These concepts span multiple levels of abstraction:

- **Concrete letters:** The 26 lowercase letters (a–z), representing the atomic units of the problem domain
- **Numeric lengths:** The numbers 1–5, used to describe group sizes
- **Positional concepts:** leftmost, rightmost, first, last, middle
- **Relational concepts:** successor, predecessor, sameness
- **Category concepts:** letterCategory, bondCategory, groupCategory
- **Meta-concepts:** opposite, identity

The slipnet contains 202 directed links connecting these nodes. When converted to an undirected graph, this yields 104 unique edges after removing directional duplicates.

## 1.2 Conceptual Depth

Each slipnet node has a *conceptual depth* parameter, a numeric value between 10 and 90 representing its level of abstraction. Hofstadter and Mitchell intended this parameter to capture the “deepness” of a concept—how far removed it is from surface-level, perceptual features:

- Letter nodes (a–z): depth = 10 (most concrete)
- `letter`: depth = 20
- `letterCategory`, numbers 1–5: depth = 30
- `leftmost`, `rightmost`, `middle`: depth = 40
- `predecessor`, `successor`: depth = 50
- `first`, `last`, `length`: depth = 60
- `stringPositionCategory`, `directionCategory`: depth = 70
- `sameness`, `samenessGroup`, `group`: depth = 80
- `opposite`, `identity`, `bondFacet`, `objectCategory`: depth = 90 (most abstract)

The conceptual depth influences Copycat’s behavior in several ways: it affects activation spreading dynamics, it modulates the system’s preference for discovering “deep” versus “shallow” analogies, and it contributes to the calculation of conceptual similarity between structures.

## 1.3 Research Question

A natural hypothesis is that deeper (more abstract) concepts should be topologically farther from concrete letters in the network. After all, if conceptual depth represents abstraction level, one might expect that reaching abstract concepts requires traversing more edges from the concrete letter nodes. We test this hypothesis using hop count—the minimum number of edges to traverse—as an Erdős number-style metric, with letters serving as the “center” analogous to Erdős himself.

## 2 Methods

### 2.1 Data Extraction

The slipnet structure was extracted from the original Copycat Python implementation and serialized to JSON format. The extraction preserved all 59 nodes with their attributes (name, conceptual depth, intrinsic link length) and all 202 directed links with their attributes (source, destination, fixed length, type, optional label).

### 2.2 Graph Construction

We constructed an undirected graph  $G = (V, E)$  from the slipnet using the NetworkX library. Each node in the slipnet became a vertex in  $G$ , and each directed link became an undirected edge. When multiple directed links existed between the same pair of nodes (e.g., both  $a \rightarrow b$  and  $b \rightarrow a$ ), they were collapsed into a single undirected edge. This yielded  $|V| = 59$  vertices and  $|E| = 104$  edges.

### 2.3 Hop Count Computation

For each non-letter node  $v \in V$ , we computed the minimum number of edges to reach any letter node  $\ell \in L$  where  $L = \{a, b, c, \dots, z\}$ :

$$\text{hops}(v) = \min_{\ell \in L} |P(v, \ell)| - 1 \quad (1)$$

where  $P(v, \ell)$  is the shortest path (sequence of vertices) from  $v$  to  $\ell$ . The subtraction of 1 converts path length (number of vertices) to hop count (number of edges).

This metric is analogous to an Erdős number, with the 26 letter nodes collectively playing the role of Erdős. A node with hop count 1 is directly connected to at least one letter; a node with hop count 2 is connected to a node that is connected to a letter; and so on.

### 2.4 Handling Unreachable Nodes

Five nodes in the slipnet are topologically disconnected from the letter subgraph. Rather than exclude these nodes from analysis, we assigned them a penalty distance:

$$\text{hops}_{\text{unreachable}} = 2 \times \max_{v \in V_{\text{reachable}}} \text{hops}(v) \quad (2)$$

With the maximum observed hop count among reachable nodes being 4, unreachable nodes were assigned hops = 8. This approach ensures all 33 non-letter nodes are included in the analysis while appropriately penalizing disconnected nodes.

## 2.5 Statistical Analysis

We computed both Pearson’s correlation coefficient  $r$  (measuring linear relationship) and Spearman’s rank correlation  $\rho$  (measuring monotonic relationship) between conceptual depth and hop count. Statistical significance was assessed at  $\alpha = 0.05$ .

Linear regression was performed to characterize any trend:

$$\text{hops} = \beta_0 + \beta_1 \times \text{depth} + \epsilon \quad (3)$$

The coefficient of determination  $R^2$  was computed to quantify the proportion of variance in hop count explained by conceptual depth.

## 3 Results

### 3.1 Network Connectivity

Of the 59 total nodes, 26 are letter nodes (which have hop count 0 by definition) and 33 are non-letter concept nodes. Among these 33 nodes, 28 are reachable from at least one letter and 5 are disconnected from the letter subgraph. The five disconnected nodes are:

- **identity** (depth = 90)
- **opposite** (depth = 90)
- **objectCategory** (depth = 90)
- **group** (depth = 80)
- **letter** (depth = 20)

### 3.2 Hop Distribution

Table 1 shows the distribution of hop counts among all 33 non-letter nodes.

Table 1: Distribution of minimum hops to letter nodes

Hops	Count	Example Nodes
1	3	<b>letterCategory</b> , <b>first</b> , <b>last</b>
2	6	<b>leftmost</b> , <b>length</b> , <b>bondFacet</b>
3	12	Numbers 1–5, <b>sameness</b> , <b>groupCategory</b>
4	7	<b>bondCategory</b> , <b>predecessor</b> , <b>middle</b>
8	5	<b>identity</b> , <b>opposite</b> , <b>letter</b> (unreachable)

The distribution shows most nodes (28 of 33) within 4 hops of a letter, with 5 nodes forming a disconnected cluster.

### 3.3 Descriptive Statistics

Table 2 summarizes the distributions of conceptual depth and hop count.

Table 2: Descriptive statistics for analyzed nodes (n=33)

Statistic	Depth	Hops
Minimum	20	1
Maximum	90	8
Mean	55.76	3.61
Std. Dev.	21.89	2.04

### 3.4 Correlation Analysis

The correlation analysis yielded the following results:

- Pearson correlation:  $r = 0.281$ ,  $p = 0.113$
- Spearman correlation:  $\rho = 0.141$ ,  $p = 0.433$
- Coefficient of determination:  $R^2 = 0.079$

- Linear regression:  $\text{hops} = 0.026 \times \text{depth} + 2.14$

Neither correlation coefficient approaches statistical significance. The p-values of 0.113 and 0.433 are above the 0.05 threshold. The  $R^2$  of 0.079 indicates that conceptual depth explains only 7.9% of the variance in hop count—a weak effect at best.

The regression slope of 0.026 suggests that a 10-point increase in conceptual depth predicts only a 0.26 increase in hop count—modest compared to the 2.04 standard deviation of hops.

### 3.5 Visualization

Figure 1 displays the scatter plot of conceptual depth versus minimum hops. Unreachable nodes (hops=8) are shown in red. The wide spread of depths at each hop level and the weak regression line visually confirm the absence of any strong relationship.

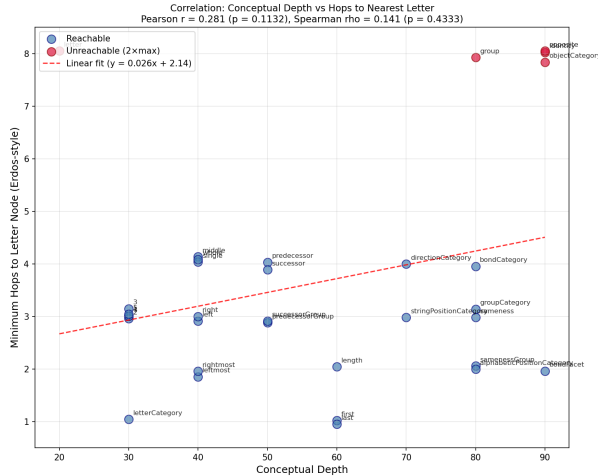


Figure 1: Scatter plot of conceptual depth versus minimum hops to nearest letter node. Reachable nodes (blue) and unreachable nodes (red, assigned hops=2 × 4 = 8) are distinguished. Points are jittered vertically for visibility. The dashed line shows the linear regression fit.

### 3.6 Counterexamples

The data reveal striking counterexamples to any depth-distance relationship:

1. **High depth, few hops:** `bondFacet` (depth=90, the maximum) is only 2 hops from a letter. Similarly, `samenessGroup` and `alphabeticPositionCategory` (both depth=80) are also just 2 hops away.
2. **Low depth, many hops:** The `letter` node (depth=20) is completely disconnected from actual letters despite being the object-type concept for them. The number nodes 1–5 (depth=30) all require 3 hops to reach a letter.
3. **Same depth, different hops:** At depth=90, `bondFacet` needs only 2 hops while `identity`, `opposite`, and `objectCategory` are completely unreachable—a dramatic difference.
4. **Same hops, different depths:** Nodes at 2 hops have depths ranging from 40 (`leftmost`) to 90 (`bondFacet`)—the full 50-point range.
5. **Unreachable nodes span depths:** The 5 disconnected nodes have depths of 20, 80, and 90—covering most of the depth range despite all being topologically equivalent (infinitely far from letters).

## 4 Discussion

### 4.1 Orthogonal Design Dimensions

The weak, non-significant correlation ( $r = 0.281$ ,  $p = 0.113$ ) demonstrates that conceptual depth and network topology were designed as largely independent dimensions. This orthogonality is architecturally meaningful:

1. **Network topology** determines which concepts can activate each other through spreading activation. Two nodes connected by an edge can directly influence each other’s activation levels during reasoning.
2. **Conceptual depth** modulates how the system values discoveries at different abstraction levels. Deeper concepts, when activated, contribute more to the system’s sense of having found a “good” analogy.

By keeping these dimensions independent, the slipnet can connect concepts that need to interact (regardless of depth) while separately encoding their semantic abstraction level.

## 4.2 The Disconnected Cluster

The five disconnected nodes form a coherent subsystem:

- **identity** and **opposite**: These exist primarily as labels on slip links, not as endpoints in the graph. They track activation for meta-level relationship concepts.
- **letter**, **group**, **objectCategory**: These form an isolated cluster representing the object-type hierarchy. They classify workspace objects but don't connect to the letter-category network.

Notably, the **letter** concept (depth=20, relatively concrete) is disconnected while **letterCategory** (depth=30) is directly connected to all 26 letters. This distinction between “letter-as-type” and “letter-as-category” further illustrates how topology and depth serve different purposes.

## 4.3 Hub Structure

Analysis of the shortest paths reveals that routes to letters converge on gateway nodes:

- **first** → **a**: Property link providing direct access
- **last** → **z**: Property link providing direct access
- **letterCategory** → any letter: Instance links to all 26 letters

The **letterCategory** node is particularly important, serving as a central hub. This makes it the primary gateway between abstract concepts and concrete letters, explaining why many paths route through it.

## 4.4 Implications

Our findings have implications for understanding and extending the Copycat architecture:

1. **For analysis**: Attempting to infer conceptual depth from topology—or vice versa—would be misguided. They encode different information.
2. **For extensions**: New concepts added to the slipnet can be placed topologically based on needed associations, with depth set independently based on abstraction level.
3. **For interpretation**: The slipnet's representational power comes from having multiple orthogonal dimensions, not from a single unified hierarchy.

## 4.5 Limitations

Several limitations should be noted:

1. **Sample size**: With 33 nodes, statistical power is limited, though this represents the complete population of non-letter nodes.
2. **Penalty assignment**: The choice of  $2 \times \max(\text{hops})$  for unreachable nodes is somewhat arbitrary. However, alternative penalty values (e.g.,  $3 \times \max$  or  $\infty$ ) would likely strengthen our conclusion.
3. **Undirected assumption**: We treated edges as undirected. Analysis of directed paths might differ.
4. **Single metric**: Hop count is one of many possible graph metrics. Centrality measures or spectral properties might reveal different patterns.

## 5 Conclusion

There is no statistically significant relationship between conceptual depth and hop distance to letter nodes in the Copycat slipnet. With Pearson  $r = 0.281$  ( $p = 0.113$ ), Spearman  $\rho = 0.141$  ( $p = 0.433$ ), and  $R^2 = 0.079$ , conceptual depth

explains less than 8% of the variance in topological distance—and this weak positive trend fails to reach significance.

This finding supports the view that the slipnet employs two orthogonal representational dimensions: network topology (governing associative access and activation flow) and conceptual depth (governing abstraction-level preferences in reasoning). This separation allows independent tuning of each dimension and may contribute to the slipnet’s representational flexibility.

## Data Availability

All analysis scripts and data are available in the `slipnet_analysis/` directory:

- `slipnet.json`: Complete network with computed paths
- `compute_letter_paths.py`: Hop computation script
- `plot_depth_distance_correlation.py`: Statistical analysis and plotting
- `compute_stats.py`: Detailed statistics computation

## A Complete Data

Table 3 presents all 33 analyzed nodes sorted by hop count and depth.

Table 3: All analyzed nodes sorted by hop count, then depth

Node	Depth	Hops	Reachable
letterCategory	30	1	Yes
first	60	1	Yes
last	60	1	Yes
leftmost	40	2	Yes
rightmost	40	2	Yes
length	60	2	Yes
samenessGroup	80	2	Yes
alphabeticPositionCategory	80	2	Yes
bondFacet	90	2	Yes
1	30	3	Yes
2	30	3	Yes
3	30	3	Yes
4	30	3	Yes
5	30	3	Yes
left	40	3	Yes
right	40	3	Yes
predecessorGroup	50	3	Yes
successorGroup	50	3	Yes
stringPositionCategory	70	3	Yes
sameness	80	3	Yes
groupCategory	80	3	Yes
middle	40	4	Yes
single	40	4	Yes
whole	40	4	Yes
predecessor	50	4	Yes
successor	50	4	Yes
directionCategory	70	4	Yes
bondCategory	80	4	Yes
letter	20	8	No
group	80	8	No
identity	90	8	No
opposite	90	8	No
objectCategory	90	8	No

## B Link Type Distribution

The slipnet contains five distinct types of directed links, summarized in Table 4.

Table 4: Slipnet link type distribution

Type	Count	Purpose
nonSlip	83	Lateral associations that don't allow conceptual slippage
category	51	Upward hierarchy (instance to category)
instance	50	Downward hierarchy (category to instance)
slip	16	Links allowing conceptual slippage
property	2	Intrinsic attributes ( $\mathbf{a} \rightarrow \mathbf{first}$ , $\mathbf{z} \rightarrow \mathbf{last}$ )

## References

- [1] Mitchell, M. (1993). *Analogy-Making as Perception: A Computer Model*. MIT Press.
- [2] Hofstadter, D. R., & FARG. (1995). *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. Basic Books.