A Mathematical Framework for Evaluating the Creativity of Ideas

Samuel Schapiro*
Spiral Works
sam@spiralworks.ai

Royce Moon †
Spiral Works
royce@spiralworks.ai

Aishik Sanyal
Spiral Works
aishik@spiralworks.ai

Abstract

Creative reasoning is a crucial aspect of problem solving and a significant driver of human progress. While prevailing theories of creativity in psychology and philosophy are abundant (Simonton, 2023; Thagard, 2012; Boden, 2004; Koestler, 1964; Güss et al., 2021), they cannot be used to precisely quantify the creativity of individual ideas. To overcome this issue, we synthesize past philosophical and psychological studies of creativity into a mathematical framework for evaluating the creativity of scientific and industrial ideas, intended to serve broadly as a foundation for future research in creativity. We view our theoretical contributions as a critical first step towards consolidating our understanding of creativity in the short term, enhancing the creativity of generative models in the medium term, and automating creative discovery in science and industry in the long term.

1 Mathematical Framework

Definition 1.1 (Domain) *Let $\mathcal{C}$ refer to the set of all concepts and $\mathcal{R}$ to the set of all logical and semantic relations in the entirety of human knowledge, and let $\mathcal{D}$ refer to a particular scientific or industrial field (e.g., chemistry, marketing, mathematics, etc.) or subfield (e.g., drug discovery, algorithm design). A **scientific or industrial domain** D is a triple $(\mathcal{M}_\mathcal{D}, \mathcal{I}_\mathcal{D}, \mathcal{T}_\mathcal{D})$, where*

1. *$\mathcal{M}_\mathcal{D} := (\mathcal{C}_\mathcal{D}, Dis_\mathcal{D})$ is a **conceptual reasoning space** consisting of (i) a subset of concepts $\mathcal{C}_\mathcal{D} \subset \mathcal{C}$ relevant to $\mathcal{D}$ and (ii) a function of semantic distance, computed over $\mathcal{D}$, between concepts $Dis_\mathcal{D} : \mathcal{C}_\mathcal{D} \times \mathcal{C}_\mathcal{D} \to [0, 1]$ (e.g., cosine similarity, LSA (Deerwester et al., 1990)).*[3]

2. *$\mathcal{I}^\alpha_\mathcal{D} := \{I = ((C_1, R_1, C_2, R_2, \ldots, C_n),\ g) \mid (\forall i \in \{1, \ldots, n\})(C_i \in \mathcal{C}_\mathcal{D}, R_i \in \mathcal{R}),$ $g \in \mathcal{G}_\mathcal{D},\ U_\mathcal{D}(I) \geq \alpha,\ n \geq 2\}$ is the **set of truths, facts, theorems, laws, etc. belonging to** $\mathcal{D}$ (Simonton, 2004), i.e., the set of existing ideas (defined in Definition 1.2) with utility (defined in Definition 1.3) above some $\alpha > 0$; and*

3. *$\mathcal{T}_\mathcal{D} := (\mathcal{G}_\mathcal{D}, \mathcal{P}_\mathcal{D})$ is a tree of depth $d \in \mathbb{N}$ representing the hierarchical goal structure of $\mathcal{D}$ (Chung et al., 1989), with node set $\mathcal{G}_\mathcal{D}$, edge set $\mathcal{P}_\mathcal{D}$, and $f : \mathcal{G}_\mathcal{D} \to \mathbb{N}$ giving the 0-indexed depth of each node. The root node is given by $g_0 \in \mathcal{G}_\mathcal{D}$ and the tree is constructed recursively as follows: for all $i < d$, g' is a child node of g_i if solving g' is necessary to solving g_i.*

Definition 1.2 (Idea) *An idea $I := ((C_1, R_1, C_2, R_2, \ldots, C_n), g)$ is a tuple of (i) a chain of thought $(C_1, R_1, C_2, R_2, \ldots, C_n)$ formed by a **conceptual path** $(C_1, \ldots, C_n) \in \mathcal{C}^n$ joined pairwise by a **relational reasoning path** $(R_1, \ldots, R_{n-1}) \in \mathcal{R}^{n-1}$ for some n, and (ii) a goal $g \in \mathcal{G}_\mathcal{D}$ it solves.*

---

*Also affiliated with the University of Illinois, Urbana-Champaign; part of this work done while visiting the Simons Institute for the Theory of Computing, Berkeley.

†Also affiliated with the University of Michigan.

[3]We assume the semantic dissimilarity $\mathrm{Dis}_\mathcal{D}(C_i, C_j) = 1$ for all concept pairs $C_i, C_j \notin \mathcal{C}_\mathcal{D}$.

Submitted to the NeurIPS 2024 Workshop on Creativity & Generative AI

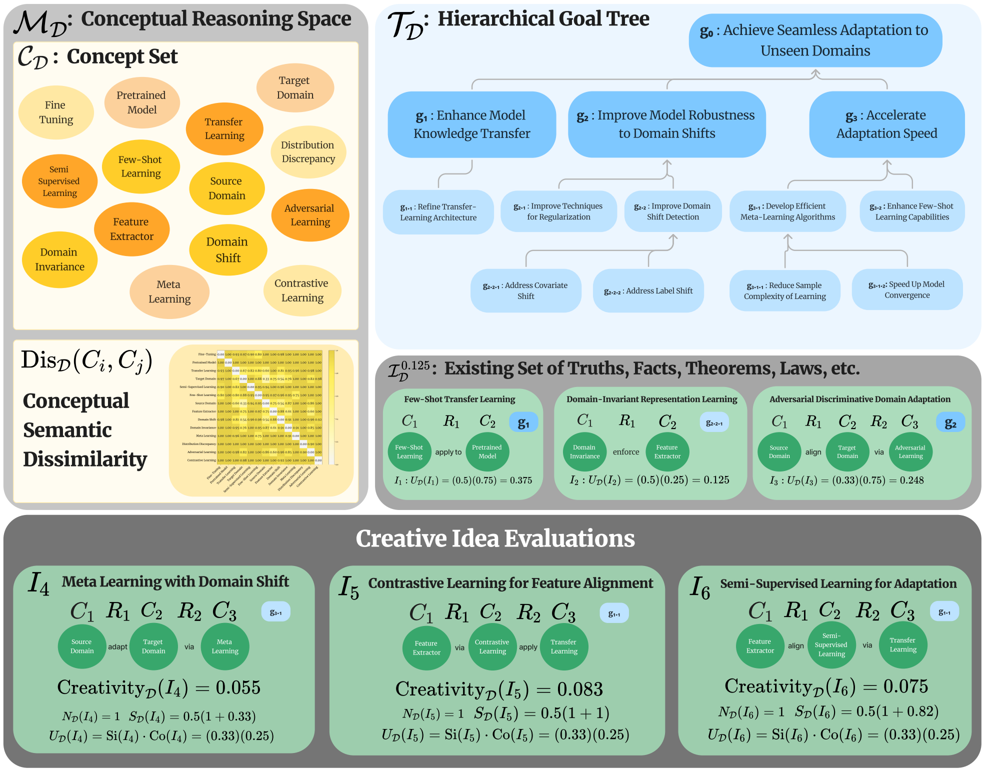

Figure 1: An example of Definitions 1.1, 1.2, & 1.3 for **domain generalization in machine learning**.

Definition 1.3 (**D-Creativity**) *The D-Creativity of an idea I is the product of its novelty, utility, and surprise (Simonton, 2004, 2023). That is, $\text{Creativity}_D(I) := N_D(I) \cdot U_D(I) \cdot S_D(I)$, where*

1. *$N_D(I) := \mathbb{I}[I \notin \mathcal{I}_\mathcal{D}^\alpha] \in \{0, 1\}$ is the* ***novelty*** *of I, indicating whether or not the idea already exists in the set of truths, facts, theorems, laws, etc. belonging to $\mathcal{D}$;*
2. *$U_D(I) := Si(I) \cdot Co_D(I)$ is the* ***utility*** *of I, defined as the product of its simplicity and consilience (Thagard, 1988):*
   (a) *$Si(I) := 1/n$ is the* ***simplicity*** *of I, the inverse of the length of its conceptual path;*
   (b) *$Co_D(I) := (d - f(g))/d$ is the* ***consilience*** *of the idea I, where d is the depth of $\mathcal{T}_\mathcal{D}$ and $f(g)$ is the depth of goal g, e.g., a depth 1 goal in a depth 5 tree evaluates to $4/5$.*
3. *$S_D(I) := \frac{1}{n-1}\sum_{i=1}^{n} Dis_D(C_i, C_{i+1})$ is the* ***surprise*** *of I, defined as the average pairwise semantic dissimilarity of its conceptual path, reflecting a departure from $\mathcal{D}$'s a priori beliefs*[4]

In Figure 1, with **domain generalization in machine learning** as the choice of $\mathcal{D}$, we instantiate the framework in Definitions 1.1 and 1.2 to evaluate the D-creativity of three ideas using Definition 1.3

Future use cases for this framework are numerous, including *(i) as a post-processing, fine-tuning, or prompt engineering template* for pre-trained models to guide creative outputs in a theoretically supported way, *(ii) as a blueprint for future automated creativity systems* that can automatically construct domains (Definition 1.1) and produce candidate ideas (Definition 1.2), *(iii) as a starting point for future mathematical study of creativity*, and more. Limitations of this framework include its use of *(i) binary-valued novelty* when novelty may actually occur in degrees, *(ii) that ideas are presupposed to solve goals* without having been verified for logical consistency or evaluated as to how well they solve goals, and that *(iii) ideas are assumed to only solve one goal*, when in reality, ideas may solve multiple goals simultaneously.

[4]This is inspired by forward flow, proposed in Gray et al. (2019) to measure the capacity for human creativity as the ability to make distant associations (Simonton, 2004).

References

Boden Margaret A. The creative mind: Myths and mechanisms. 2004.

Chung Donald T, Modarres Mohammad, Hunt R Niall M. GOTRES: an expert system for fault detection and analysis // Reliability Engineering & System Safety. 1989. 24, 2. 113–137.

Deerwester Scott, Dumais Susan T, Furnas George W, Landauer Thomas K, Harshman Richard. Indexing by latent semantic analysis // Journal of the American society for information science. 1990. 41, 6. 391–407.

Gray Kurt, Anderson Stephen, Chen Eric Evan, Kelly John Michael, Christian Michael S, Patrick John, Huang Laura, Kenett Yoed N, Lewis Kevin. "Forward flow": A new measure to quantify free thought and predict creativity. // American Psychologist. 2019. 74, 5. 539.

Güss C Dominik, Ahmed Sarah, Dörner Dietrich. From da Vinci's flying machines to a theory of the creative process // Perspectives on Psychological Science. 2021. 16, 6. 1184–1197.

Hupkes Dieuwke, Giulianelli Mario, Dankers Verna, Artetxe Mikel, Elazar Yanai, Pimentel Tiago, Christodoulopoulos Christos, Lasri Karim, Saphra Naomi, Sinclair Arabella, others . A taxonomy and review of generalization research in NLP // Nature Machine Intelligence. 2023. 5, 10. 1161–1174.

Khoee Arsham Gholamzadeh, Yu Yinan, Feldt Robert. Domain generalization through meta-learning: a survey // Artificial Intelligence Review. 2024. 57, 10. 285.

Koestler Arthur. The act of creation. 1964.

Lee Hung yi, Li Shang-Wen, Vu Ngoc Thang. Meta Learning for Natural Language Processing: A Survey. 2022.

Liu Jiashuo, Shen Zheyan, He Yue, Zhang Xingxuan, Xu Renzhe, Yu Han, Cui Peng. Towards Out-Of-Distribution Generalization: A Survey. 2023.

Ramponi Alan, Plank Barbara. Neural Unsupervised Domain Adaptation in NLP—A Survey. 2020.

Rohlfs Chris. Generalization in Neural Networks: A Broad Survey. 2024.

Sheth Paras, Moraffah Raha, Candan K. Selccuk, Raglin Adrienne, Liu Huan. Domain Generalization – A Causal Perspective. 2022.

Simonton Dean Keith. Creativity in science: Chance, logic, genius, and zeitgeist. 2004.

Simonton Dean Keith. The blind-variation and selective-retention theory of creativity: Recent developments and current status of BVSR // Creativity Research Journal. 2023. 35, 3. 304–323.

Thagard Paul. Computational Philosophy of Science. 1988.

Thagard Paul. Creative Combination of Representations: Scientific Discovery and Technological Invention // Psychology of Science: Implicit and Explicit Processes. 07 2012.

Wang Jindong, Lan Cuiling, Liu Chang, Ouyang Yidong, Qin Tao, Lu Wang, Chen Yiqiang, Zeng Wenjun, Yu Philip S. Generalizing to Unseen Domains: A Survey on Domain Generalization. 2022.

Wang Jindong, Li Haoliang, Pan Sinno, Xie Xing. A Tutorial on Domain Generalization // Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. 2023. 1236–1239.

Zhou Kaiyang, Liu Ziwei, Qiao Yu, Xiang Tao, Loy Chen Change. Domain Generalization: A Survey // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022. 1–20.

A Appendix

A.1 Example Details

In Figure 1, we instantiate the framework defined in Definitions 1.1 and 1.2 to evaluate the creativity of three candidate ideas in the chosen subfield **domain generalization in machine learning** according to Definition 1.3. In order to compute the semantic dissimilarity function $\text{Dis}_{\mathcal{D}} : \mathcal{C}_{\mathcal{D}} \times \mathcal{C}_{\mathcal{D}} \to [0, 1]$, we use a variant of latent semantic analysis defined in Gray et al. (2019), where we count the co-occurrence frequency of concept pairs over a corpus of text, consisting of a concatenated set of 10 survey papers in domain generalization (Ramponi, Plank, 2020; Zhou et al., 2022; Wang et al., 2022; Liu et al., 2023; Lee yi et al., 2022; Rohlfs, 2024; Sheth et al., 2022; Wang et al., 2023; Khoee et al., 2024; Hupkes et al., 2023).

Semantic Dissimilarity Measure Frequency of co-occurrence is a basic measure of semantic dissimilarity that only serves to illustrate the basic example we construct in this paper. Future studies using this framework as a starting point for obtaining creative outputs in a theoretically supported way should investigate more elaborate semantic dissimilarity measures.

Out of Domain Knowledge Although the examples portrayed in Figure 1 consist solely of concepts in the concept set $\mathcal{C}_{\mathcal{D}}$ for the chosen example domain, in practice, novelty can be achieved by incorporating concepts in $\mathcal{C} \setminus \mathcal{C}_{\mathcal{D}}$ (out of domain knowledge) into ideas intended to solve goals belonging to the goal tree $\mathcal{T}_{\mathcal{D}}$ in domain $\mathcal{D}$.