



**Open Access**  
**Research Article**  
Received 24th February 2026  
Accepted 1st March 2026  
**MID: None**

# RUBBISH

## 语言概率学——以意大利面拌 42 号混凝土为例

水何泠泠<sup>1,\*†</sup>, Deepseek<sup>2,\*†</sup>

### Abstract

随着大语言模型的迅猛发展，以数学概率视角审视语言现象已成为现实可能。本文从网络热梗“意大利面拌 42 号混凝土”切入，运用概率学方法分析这一低概率表达为何能成为广泛传播的流行语，揭示其背后的语言概率学原理，并展望这一新兴交叉学科的研究前景与潜在应用。最后，本研究及其局限，请勿太过认真。

**Keywords:** Classification 1; Classifications 2; Classifications 3

### 1 Introduction

“意大利面拌 42 号混凝土”出自抖音博主“延边刺客”。原话将建筑、机械、生物等领域的术语暴力嫁接，逻辑不通却一本正经，博主因此被称为“知识搅拌机”。2024 年春节期间，该梗二次爆发，成为年轻人应对亲戚盘问的“发疯文学”——只需正色背诵这段台词，即可用荒诞建立社交边界，迅速出圈。这句话本意是拼接风马牛不相及的术语以展现精神失常。然而，“意大利面”和“42 号混凝土”真的没有关系吗？

意大利面：属于食品科学和农业领域。主要原

料是硬质小麦粗面粉和水，有的会加鸡蛋。它的核心功能是被人体消化吸收以提供能量和营养。

42 号混凝土：属于土木工程和材料科学领域。混凝土通常由胶凝材料（水泥）、骨料（砂、石）和水按一定比例混合而成。所谓的“42 号”通常指水泥的强度等级（即 42.5MPa）。它的核心功能是作为结构材料承受巨大的压力和荷载。

在绝大多数日常对话、文章、书籍中，“意大利面”和“42 号混凝土”是毫不相干的两个概念。它们出现在同一个句子中的概率已经极低，而连续出现的概率更是微乎其微。从概率学上讲，假设汉语词汇总量约为 105105 量级，如果每个词随机出现，

那么给定前词后，下一个词是特定词的概率约为 $10^{-5}$ 。但实际语言中词与词有强烈的语法和语义约束，这种完全不相关的搭配概率会比随机还低，因为语言模型会给予极低概率。

那么，意大利面拌 42 号混凝土为什么会产生呢？

## 2 Methodology

意大利面与 42 号混凝土之间的直接联系并不深，甚至及其微小，但是‘意大利面拌 42 号混凝土’则不一样，‘拌’这个字在其中起到了桥梁作用。

在正常的语言模型中，我们的大脑会下意识地选择最大似然估计——即最可能出现的下一个词。说到‘意大利面’，后面大概率可能是番茄酱，说到‘混凝土’，前面大概率可能是搅拌机。而‘意大利面拌 42 号混凝土’这个短语，其实是以‘拌’这个字为桥梁，将前面与后面两个跨领域术语衔接到一起。

### 2.1 意大利面后面接‘拌’的概率

在汉语中，提到“意大利面”时，后面常跟的动词有“煮”、“吃”、“做”、“炒”、“拌”等，也可能跟名词如“的”、“酱”等。

“拌”是制作意大利面时的一个常见步骤（尤其是与酱料混合时），但相对于“煮”和“吃”来说，它出现的频率稍低。在菜谱或日常对话中，大约有 5%~15% 的“意大利面”后面会直接跟“拌”（例如“意大利面拌肉酱”）。

考虑到语言多样性，我们取一个中间值，

大约在 0.05 到 0.1 之间。

### 2.2 ‘拌’后面接‘42号混凝土’的概率

“拌”是一个烹饪/混合动词，后面通常跟可食用的食材，如“黄瓜”、“面”、“沙拉”、“凉菜”、“馅”等。在建筑领域，“拌”也可以跟“水泥”、“砂浆”等，但“42 号混凝土”是特定标号的建筑材料，极少出现在日常语境中。

在绝大多数文本中，“拌”后面出现“42 号混凝土”的概率几乎为零。即使考虑所有可能的搭配，这个概率也远小于  $10^{-6}$ （百万分之一），因为“42 号混凝土”本身就是一个低频词，且与“拌”的语义场完全不搭。

在正常语言模型中，这个概率可以认为是  $<10^{-6}$ ，甚至更小。

### 2.3 拌’后接‘混凝土’的概率

对比“拌”后面跟“42 号混凝土”的概率（我们之前估计为  $<10^{-6}$ ），去掉“42 号”这个限定后，可能的宾语范围扩大到“混凝土”、“水泥”、“砂浆”等。其中，“混凝土”作为一个更通用的词，出现频率比“42 号混凝土”高得多，约为  $10^{-4} \sim 10^{-5}$ ，但整体仍然极小。然而在建筑领域，‘搅拌混凝土’本身却是高频词。

### 2.3.4 联合概率

短语“意大利面拌 42 号混凝土”在正常语言中出现的概率约为：

---

<sup>1</sup>First Institution and <sup>2</sup>Second Institution

\*abc@uni.edu; alphanbeta@lab.edu

†Contributed equally.

$P(\text{意大利面拌 42 号混凝土})$

$$\begin{aligned} &\approx P(\text{意大利面}) \times P(\text{拌} \\ &| \text{意大利面}) \times P(\text{42 号混凝土} \\ &| \text{拌}) \approx P(\text{意大利面}) \times (0.05 \\ &\sim 0.1) \times (< 10^{-6}) \end{aligned}$$

### 3 ‘意大利面拌 42 号混凝土’ 概率分析

正是因为它在正常语言中几乎不可能出现，所以当有人刻意将它创造出来时，会产生强烈的新奇感和荒谬感。这种创造不是随机碰撞，而是创作者（延边刺客）故意违背语言常规，将两个毫无关联的领域通过一个常见的动词“拌”强行连接，制造出“一本正经地胡说八道”的喜剧效果。在网络时代，这种打破常规的创意一旦出现，就会被迅速传播和模仿，从而成为一个流行的梗。

## 4 Conclusion

从上面的分析中，我看到了一个新的学问——语言概率学的前景，语言本身是一门无法用数学衡量的文明基石，但是在本质是计算语言概率模型的大模型的应用下，每一句话都可以用数字来衡量，小到创作网络热梗，大到分析未知文明，语言概率学或将成为一门用概率论、统计学和信息论的视角，研究语言的生成、理解、演化与创新的交叉学科。

## 致谢

感谢 deepseek 提供的概率计算与分析，如有任何疑问，那就是我的错，写论文发疯产物，请勿认真

## Reference

[1] deepseek