arXiv:2303.05616v3 [cs.CC] 6 Sep 2023

如何压缩解

Samuel Epstein*

2023年9月8日

摘要

我们利用去随机化，给出了图着色问题的解的压缩规模上界。一般来说，如果某个组合问题的解以高概率存在，且该概率是简单的，那么该问题就存在一个简单解。否则，该问题实例与停机问题具有较高的互信息。

1 引言

在数学中，概率方法是一种构造性方法，用于证明某类数学对象的存在性。该方法由保罗·厄尔德什首创，其做法是从某个类别中随机选取对象，并证明某类对象出现的概率不为零。因此，可以保证这类对象必然存在。有关概率方法的更多信息，请参阅 ¨读者可参考 [AS04]。近期研究结果表明，概率方法与数学对象编码的压缩大小之间存在紧密联系，即它们的柯尔莫哥洛夫复杂度，$\mathbf{K}$：

如果可以用概率方法证明某个对象存在，那么也可以证明其柯尔莫哥洛夫复杂度的上界。

如果存在一种简单的概率分布，使得某类数学对象以较大的概率出现，那么就存在一个简单的该类对象，也就是说，它的柯尔莫哥洛夫复杂度较低。更形式化地说，如果对象 x 随机出现的概率至少为 $P\ p$，那么

$$\mathbf{K}(x) <^{\log} \mathbf{K}(P) - \log p + \epsilon.$$

该 ϵ 项表示对整个数学结构进行编码后所得信息中包含的停机序列信息量；显然，这部分信息可以视为可忽略不计，除非是在一些特殊情形下。

这个不等式是通过应用 EL 定理 [Lev16，Eps19]得出的。通过概率方法对对象的柯尔莫哥洛夫复杂度给出界，称为去随机化。在 [Eps22a], 中，去随机化被应用于 22 个示例，其中包括若干博弈。在 [Eps22b], 中，去随机化被用于证明经典信道容量与码本复杂度之间的权衡。此外，还引入了时间资源有界去随机化。

我建议有意转向算法信息论的硕士生或研究人员将去随机化作为一个研究方向。这是因为大部分技术工作都集中在去随机化所应用的领域。

本文展示了一个典型的去随机化示例，即图顶点着色。证明需要应用 EL 定理、应用守恒定理，并进行一些直接的概率论论证。

*JP 理论组。samepst@jptheorygroup.org

2 结果

函数 K($x|y$) 是条件前缀柯尔莫哥洛夫复杂度。算法概率 $\mathbf{m}(x) = \{2^{-\|p\|} : U(p) = x\}$，其中 U 是通用图灵机。函数 $\mathbf{m}$ 是通用的，也就是说，对于任意在 $\{0,1\}^*$上的可计算概率 P， $O(1)$ $\mathbf{m}(x) > 2^{-\mathbf{K}(P)}P(x)$。因此，对于集合 $D \subseteq \{0,1\}^*$，可计算概率 P， $O(1)\mathbf{m}(D) > 2^{-\mathbf{K}(P)}P(D)$。$a \in \{0,1\}^*$与停机序列 $\mathcal{H} \in \{0,1\}^\infty$之间的信息为 $\mathbf{I}(a;\mathcal{H}) = \mathbf{K}(a) - \mathbf{K}(a|\mathcal{H})$，其中 $\mathcal{H} \in \{0,1\}^\infty$是停机序列。 $<^+ f$是 $<f+O(1)$，而 $<^{\log} f$是 $<f+O(\log(f+1))$。下面的引理是一条信息非增长定律。此类定律有很多，这条定律讨论的是与停机序列之间的非对称信息。

引理 1 ([Eps22a])
对于部分可计算的 $f\colon \mathbb{N} \to \mathbb{N}$，对于所有 $a \in \mathbb{N}$，$\mathbf{I}(f(a);\mathcal{H}) <^+ \mathbf{I}(a;\mathcal{H}) + \mathbf{K}(f)$。

以下结果是 EL 定理 [Lev16， Eps19]。它最初被表述为关于学习的论断。然而，自那时以来，它已出现了若干意想不到的应用。在本文中，EL 定理用于去随机化

定理 1（EL 定理）对于有限的 $D \subset \{0,1\}^*$, $-\log \max_{x\in D} \mathbf{m}(x) <^{\log} -\log \mathbf{m}(D) + \mathbf{I}(D;\mathcal{H})$。

对于图 $G = (V,E)$，其边为无向边， k-着色是一个函数 $f\colon V \to \{1,\dots,k\}$，满足若 $(v,u) \in E$，则 $f(v) \neq f(u)$。图着色的一个例子是手机信号塔，每座塔都需要在特定频率上运行。由于干扰，彼此距离很近的手机信号塔不能使用相同的频率。因此，可以用顶点表示信号塔，用边表示干扰，而图着色则表示为信号塔分配频率。然而，如果要传输合法着色的描述，会发生什么情况？以下定理是去随机化的成果，它表明，使用更多颜色时，合法着色的最优编码长度会更短。

定理 2 对于图 $G = (V,E)$， $|V| = n$ 其最大度数为 d，存在一种 k 着色 C 满足 $2d \le k$，且 $\mathbf{K}(C) <^{\log} \mathbf{K}(n,k) + 2nd/k + \mathbf{I}((G,k);\mathcal{H})$。

证明。假设每个顶点都被随机赋予 $\{1,\dots,k\}$中的一种颜色。每个顶点都不与其他顶点发生冲突的概率为 $\ge \frac{k-d}{k}$。因此，均匀随机颜色赋值构成合法着色的概率为 $\ge \left(\frac{k-d}{k}\right)^n$。令有限集 $D \subset \{0,1\}^*$表示所有合法着色的编码。因此，存在一个简单的全可计算函数 f，它以 G 和 k 为输入，可输出 D，且 $\mathbf{K}(f) = O(1)$。令 P 表示对 G 各顶点的所有颜色赋值（包括不构成合法着色的赋值）所取的均匀概率。因此，考虑到 $2d \le k$，

$$-\log P(D) < -n \log\left(1 - \frac{d}{k}\right) \le \frac{2nd}{k}.$$

因此，根据通用半测度 $\mathbf{m}$ 的定义，

$$-\log \mathbf{m}(D) <^+ \mathbf{K}(P) + P(D)$$

因此，根据定理 1 和引理 1，存在一种着色 $C \in D$ 满足

$$\begin{aligned}\mathbf{K}(C) &<^{\log} -\log \mathbf{m}(D) + \mathbf{I}(D;\mathcal{H}) \\ &<^{\log} \mathbf{K}(P) - \log P(D) + \mathbf{I}(D;\mathcal{H}) \\ &<^{\log} \mathbf{K}(n,k) + 2nd/k + \mathbf{I}((G,k);\mathcal{H}).\end{aligned}$$

□

3 讨论

未来的工作包括寻找概率方法的实例，并对其应用去随机化。特别是，Lov´asz 局部引理， [EL], 与去随机化尤其兼容。我们给出 LLL 的第一个已证明推论，并展示它如何与三种去随机化方式兼容：一种涉及柯尔莫哥洛夫复杂度，一种涉及资源受限的柯尔莫哥洛夫复杂度，还有一种涉及博弈。

超图是由顶点 V 和超边 $E \subseteq \mathcal{P}$ 组成的一对 $J = (V, E)$（顶点为 V，超边为 $E \subseteq \mathcal{P}$）。因此，每条超边都可以连接 ≥ 2 个顶点。超图是 k-正则的，所有超边 $e \in E$ 的大小均为 $|e| = k$。2-正则超图就是简单图。超图 (V, E) 的一个有效 C-着色是一个映射f： $V \to \{1, \ldots, C\}$，其中每条超边 $e \in E$ 都不是单色的 $|\{f(v)\colon \ v \in e\}| > 1$。下面的经典结果利用 LLL 得到证明。

定理。（概率方法）设 $G = (V, E)$ 是一个 k-正则超图。如果对于每条超边 f，至多有 $2^{k-1}/e-1$ 条超边 $h \in E$ 满足 $h \cap f \neq \emptyset$，那么 G 存在一个有效的二染色。

现在我们可以使用去随机化，对 G 中最简单的这种二染色的柯尔莫哥洛夫复杂度给出界。

定理 A.（去随机化）设 $J = (V, E)$ 是一个 k-正则超图，且满足 $|E| = m$。如果对于每条超边 f，至多有 $2^{k-1}/e - 1$ 条超边 $h \in E$ 满足 $h \cap f \neq \emptyset$，那么存在 J 的一个有效二染色 x，使得

$$\mathbf{K}(x) <^{\log} \mathbf{K}(n) + 4me/2^k + \mathbf{I}(J; \mathcal{H}).$$

现在我们可以利用 [Eps22b] , 中的资源受限去随机化，为$\mathbf{K}^t(x) = \min\{p\colon \ U(p) = x$ 在 t ($\|x\|$) 步内$\}$ 的 二染色的最小时间受限柯尔莫哥洛夫复杂度给出界 J。

假设。Crypto 是指存在一个属于 DTIME($2^{O(n)}$)、且没有大小为 $2^{o(n)}$ 、含有 Σ_2^p 个门的电路的语言这一假设。

定理 B.（资源有界去随机化）假设 Crypto。设 $J_n = (V, E)$ 是一个 $k(n)$-正则超图，其中 $|V| = n$ and $|E| = m(n)$ 可在 n 中由统一的多项式时间算法计算。此外，对于 J_n 中的每条超边 f，至多存在 $2^{k(n)-1}/e - 1$ 条超边 $h \in E$ 满足$h \cap f \neq \emptyset$。那么存在一个多项式 p，以及 J_n 的一个有效二染色 x，使得

$$\mathbf{K}^p(x) < 4m(n)e/2^{k(n)} + O(\log n).$$

我们定义如下一个涉及超图的博弈，该博弈来自 [Eps23]。玩家可以访问一个顶点列表，目标是为该超图生成一个有效的二染色。我们假设，对于图中的每条超边 f，至多存在 $2^{k-1}/e - 1$ 条超边 h 满足 $f \cap h \neq \emptyset$。

博弈按如下方式进行。在第一轮中，环境将顶点数量告知玩家。玩家有 n 个顶点，每个顶点的初始颜色均为 1。在之后的每个回合中，环境都会将单色超边发送给玩家。玩家最多可以更改 k 个顶点的颜色，并将这些更改发送给环境。当玩家为图完成有效的二染色时，博弈结束。

定理 C。（博弈去随机化）对于 $k \geq 6$，存在一个玩家 p，能够在 $(1+\epsilon)n/k$ 轮内击败环境 q，且其 Kolmogorov 复杂度 K(p) $<^{\log}$ I(q; $\mathcal{H}$) $-\log \epsilon$，其中$\epsilon \in$ (0,1)。

参考资料

[AS04] N. Alon 和 J. Spencer。 概率方法。Wiley，纽约，2004年。[EL] P. Erdos 和 L. Lov´asz。关于 3- 染色超图及相关问题的问题与结果。 无限集与有限集，10:609–627。[Eps19] S. Epstein。关于集合的算法概率。 *CoRR*，abs/1907.04776，2019年。[Eps22a]S. Epstein。通过去随机化实现解压缩的 22 个示例。 *CoRR*，abs/2208.11562，2022年。[Eps22b]S. Epstein。不同资源约束下的去随机化。 *CoRR*，abs/2211.14640，2022年。[Eps23] Samuel Epstein。博弈去随机化，2023年。
https://www.jptheorygroup.org/doc/Game.p

df.

[Lev16] L. A. Levin. 元素最低复杂度的奥卡姆界。《纯粹与应用逻辑年鉴》，167(10):897–900, 2016。