

A Demigod's Number for the Rubik's Cube

Arturo Merino   

Department of Computer Science, University of Chile, Santiago, Chile

Bernardo Subercaseaux   

Carnegie Mellon University, Pittsburgh, PA, USA

Abstract

It is by now well-known that any state of the $3 \times 3 \times 3$ Rubik's Cube can be solved in at most 20 moves, a result often referred to as “*God's Number*”. However, this result took Rokicki et al. around 35 CPU years to prove and is therefore very challenging to reproduce. We provide a novel approach to obtain a worse bound of 36 moves with high confidence, but that offers two main advantages: (i) it is easy to understand, reproduce, and verify, and (ii) our main idea generalizes to bounding the diameter of other vertex-transitive graphs by at most twice its true value, hence the name “*demigod number*”. Our approach is based on the fact that, for vertex-transitive graphs, the diameter at most twice the average distance (of which we give a much simpler proof than in the literature). Then, by sampling uniformly random states and using a modern solver to obtain upper bounds on their distance, a standard concentration bound allows us to confidently state that the average distance is around 18.32 ± 0.18 , from where the diameter is at most 36.

2012 ACM Subject Classification Mathematics of computing → Paths and connectivity problems

Keywords and phrases Diameter, Rubik's Cube, Experimental mathematics

Digital Object Identifier 10.4230/LIPIcs.FUN.2026.31

Supplementary Material *Software*: <https://github.com/bsubercaseaux/RubikDemiGodNumber>
archived at `swb:1:dir:f60b4499805826762e42e46629f51691d98ec32c`

Funding *Arturo Merino*: supported by ANID FONDECYT Iniciación No 11251528.

Bernardo Subercaseaux: supported by NSF grant DMS-2434625.

1 Introduction

The $3 \times 3 \times 3$ Rubik's Cube, illustrated in Figure 1, is arguably one of the most iconic puzzles ever created, and one of the best-selling toys of all time; its beautiful balance of simplicity (it only has 6 faces, with 54 colored stickers) and complexity (it has over $4.3 \cdot 10^{19}$ possible states) has captured the attention of millions of people around the world since its invention in 1974. Naturally, such a combinatorially rich puzzle has raised a variety of interesting mathematical questions, with the most famous one being:

What is the minimum number of moves required to solve the Rubik's Cube from any starting position?

To make this question precise, we consider the *half-turn metric*, in which turning either of the 6 faces of the cube by any amount (i.e., 90° , 180° , or 270°) counts as a single move. After a series of incremental improvements detailed in Table 1, Rokicki et al. [19] proved that 20 moves are always enough to solve any Rubik's Cube in the half-turn metric, a result often referred to as “*God's Number*”. This result, however, was obtained by a mixture of mathematical ideas and extensive computation, taking around 35 CPU years. As a result, verifying the correctness of the so-called God's Number is extremely challenging, and the required computations have likely never been reproduced independently. While the result is widely believed to be true, our goal is to provide an alternative approach (providing a weaker bound) that can be easily understood and reproduced in e.g., a classroom setting.



© Arturo Merino and Bernardo Subercaseaux;
licensed under Creative Commons License CC-BY 4.0
13th International Conference on Fun with Algorithms (FUN 2026).
Editor: John Iacono; Article No. 31; pp. 31:1–31:20



Leibniz International Proceedings in Informatics
LIPIcs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

■ **Table 1** Historical bounds on the maximum number of moves required to solve the Rubik's Cube [17].

Year	Lower bound	Upper bound
1981	18	52
1990	18	42
1992	18	39
1992	18	37
1995	18	29
1995	20	29
2005	20	28
2006	20	27
2007	20	26
2008	20	25
2008	20	23
2008	20	22
2010	20	20

Summary

Our approach is based on the following observation: in any vertex-transitive graph (i.e., a graph where every vertex “looks the same”), the diameter D (i.e., the maximum distance between any two vertices) is at most twice the mean distance between vertices, m . This observation corresponds to a question posed by Alan Kaplan [21], which was answered in the more general context of homogeneous compact metrics spaces by Herman and Pakianathan [7]. Our proof, however, is elementary and self-contained, and we believe it can be a nice addition to a first course on graph theory.

We apply the previous observation to the Cayley graph of the Rubik's Cube (defined formally in Section 2, this graph has the possible states of the cube as vertices and edges between pairs of states that are one “move” away from each other) – if we knew the average distance μ between states of the Rubik's Cube, we could bound its diameter by at most twice this value. While we cannot directly compute the exact value of μ for the Rubik's Cube¹, we can provide a good estimate by sampling random pairs of states and computing an upper bound on their distance through standard Rubik's algorithms (i.e., Kociemba's *Two-Phase Algorithm* [13]). Each of these upper bounds for the distance between a pair of states is certified by a short sequence of moves, so one can trust the result without needing to trust the solving algorithm itself. Then, through simple concentration bounds, we will argue that the empirical average of $\hat{\mu} \approx 18.3189$ we obtain is *overwhelmingly likely* to be a good estimate of the true mean distance μ , and therefore an audience should reasonably trust that the diameter of the Rubik's Cube is at most 36 (since it must be an integer). As the first step, we will prove the following theorem.

► **Theorem 1.** *Given a state s of the Rubik's Cube, let $d(s)$ be the distance from s to the solved state. Let S be a set of states of the Rubik's Cube sampled uniformly at random, and let $\hat{\mu}_S = \frac{1}{|S|} \sum_{s \in S} d(s)$ be the random variable corresponding to the average distance*

¹ Recall that it has over $4 \cdot 10^{19}$ many states.

between states in S and the solved state. Then, if D denotes the diameter of the Rubik's Cube, we have

$$\Pr_S[D \geq 2\hat{\mu}_S + 0.36] < \exp\left(\frac{-|S|}{648\,534}\right).$$

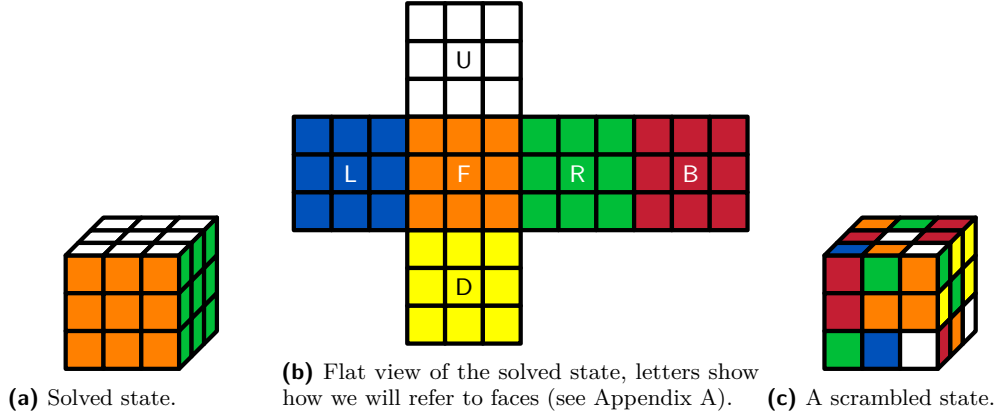
Note immediately that Theorem 1 implies that if $D > 36$ (and thus $D \geq 37$ since the diameter is an integer), then the probability of obtaining $\hat{\mu}_S \approx 18.318$ for $|S| = 10^7$ is roughly $2 \cdot 10^{-7}$. However, this experience can be observed repeatedly, which is therefore a great degree of probabilistic evidence for $D \leq 36$. Unfortunately, Theorem 1 is somewhat computationally expensive, as it requires the number of samples $|S|$ to be around 10 million if we desire a probability of error under 10^{-7} . As we show in Section 4, this would take roughly 100 hours of computation. In order to see how to reduce the required computation, let us briefly discuss how Theorem 1 is obtained. To obtain an upper bound on D , we can leverage the $D < 2\mu$ observation and look for an upper bound on μ , which we can do with a probabilistic guarantee by considering an empirical average μ_S . However, a priori it could be that the true average μ is much larger than our empirical estimate μ_S due to a small number of states that are very far and we are not likely to sample randomly – a “long tail” phenomenon. Our solution to this problem is using a “*Human's number*”, which is an unconditional modest upper bound on D . For instance, the so-called “*beginner's method*” suffices to obtain an upper bound of 205 moves. A proof sketch is provided in the appendix, but we expect readers familiar with the beginner's method to find this bound trivial, since it is far from being tight.

► **Lemma 2 (Human's Number).** *Any position of the Rubik's Cube can be solved in at most 205 moves.*

Thus, in general, our work can be interpreted as a method for transforming a “*Human's number*” into a “*Demigod number*” that is easy to trust and at most twice the real “*God's number*”. Now, going back to the problem of how to reduce the number of samples, the 648 534 constant in Theorem 1 is a consequence of the constant 205 in Lemma 2. We can improve on this by showing first that in a large percentage of the cases, we can use an upper bound of 20 instead of 205. Indeed, let us say that a state is “*far from being solved*” if it requires strictly more than 20 moves to be solved. Naturally, the “*God's number*” result corresponds to saying that no state is “*far from being solved*”, but proving this requires a great computational effort. Instead, we use a much simpler argument: if the proportion of states that are “*far from being solved*” were to be bigger than, say, 0.03%, then we would certainly expect to see a state that is “*far from being solved*” after 50 000 random samples. Yet, experimentally we do not see any such state after 500 000 samples, which makes the possibility of more than 0.03% of total states being “*far from being solved*” extremely slim. This way, we can separate μ into:

$$\left[p_{\text{far}} \cdot \sum_{s \text{ far from being solved}} d(s) \right] + \left[(1 - p_{\text{far}}) \cdot \sum_{s \text{ not far from being solved}} d(s) \right]$$

Finally, we believe that our approach can be used as an example to motivate philosophical conversations about *plausibility*, a notion explored at large by George Polya in his book “*Mathematics and Plausible Reasoning*” [16]. In a nutshell, the idea is that while certain types of reasoning do not provide full proof of a statement, they can provide a high degree of confidence in its truth [2]. This is the case, for example, with probabilistic primality tests, which may allow us to confidently assert that a number N , with hundreds of millions of digits



■ **Figure 1** Illustration of the $3 \times 3 \times 3$ Rubik's Cube.

is prime, without providing a proof in the traditional sense of the word. Then, if we proved in the traditional sense that a certain number-theoretic statement T is true provided N is prime, our confidence in the primality of N should translate into confidence in statement T .

2 Preliminaries

Let us introduce the notation and definitions required to work over the Rubik's cube mathematically. We will see the Rubik's cube as a group, a standard idea (see e.g., [4, 1]) that we make explicit nonetheless to make our work as self-contained as possible.

Let S_n denote the group of permutations of n elements under the composition operation, denoted by $\circ$. For example, $(2, 3, 1) \in S_3$ is the permutation defined by:

$$1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

and thus, for instance,

$$(2, 3, 1) \circ (1, 3, 2) = (2, 1, 3).$$

We identify each state of the Rubik's cube with a permutation of its stickers, of which there are $6 \cdot 9 = 54$ (9 per each of the 6 faces), and as moving faces in the Rubik's cube results in a permutation of its stickers, the Rubik's cube can be seen as a subgroup of S_{54} , as illustrated in Figure 2.² Furthermore, as no moves affect the relative position of the center stickers (5, 14, 23, 32, 41, 50), we can see the Rubik's cube as a subgroup of S_{48} . To complete the definition of the Rubik's cube group, we need to define the “moves”, which correspond to the following permutations (omitting the values that are not affected by the move):

$$\mathbf{R} = \begin{pmatrix} 3 & 6 & 9 & 21 & 24 & 27 & 28 & 29 & 30 & 31 & 33 & 34 & 35 & 36 & 37 & 40 & 43 & 48 & 51 & 54 \\ 43 & 40 & 37 & 3 & 6 & 9 & 30 & 33 & 36 & 29 & 35 & 28 & 31 & 34 & 54 & 51 & 48 & 21 & 24 & 27 \end{pmatrix},$$

$$\mathbf{L} = \begin{pmatrix} 1 & 4 & 7 & 10 & 11 & 12 & 13 & 15 & 16 & 17 & 18 & 19 & 22 & 25 & 39 & 42 & 45 & 46 & 49 & 52 \\ 19 & 22 & 25 & 12 & 15 & 18 & 11 & 17 & 10 & 13 & 16 & 46 & 49 & 52 & 7 & 4 & 1 & 45 & 42 & 39 \end{pmatrix},$$

² More in general, a classic theorem of Cayley states that any group is isomorphic to a subgroup of a symmetric group [10, Section 1.3].

			1	2	3			
			4	5	6			
			7	8	9			
10	11	12	19	20	21	28	29	30
13	14	15	22	23	24	31	32	33
16	17	18	25	26	27	34	35	36
			46	47	48			
			49	50	51			
			52	53	54			

■ **Figure 2** Illustration of the Rubik's cube as a subgroup of S_{54} (or even S_{48} due to the centers being static).

$$\mathbf{U} = \begin{pmatrix} 1 & 2 & 3 & 4 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 19 & 20 & 21 & 28 & 29 & 30 & 37 & 38 & 39 \\ 3 & 6 & 9 & 2 & 8 & 1 & 4 & 7 & 37 & 38 & 39 & 10 & 11 & 12 & 19 & 20 & 21 & 28 & 29 & 30 \end{pmatrix},$$

and similarly, $\mathbf{D}$, $\mathbf{F}$, and $\mathbf{B}$ can be deduced from Figure 2. The moves $\mathbf{R}'$, $\mathbf{L}'$, $\mathbf{U}'$, etc., correspond to the inverse of the moves $\mathbf{R}$, $\mathbf{L}$, $\mathbf{U}$, etc., respectively. A sequence of moves is simply a composition of moves, omitting the composition symbol, e.g., $\mathbf{R U}2$ corresponds to the permutation $\mathbf{R} \circ \mathbf{U} \circ \mathbf{U}$.

Let $\mathcal{R}$, the Rubik's cube group, be the subgroup of S_{54} generated by the standard moves, i.e., $\mathcal{R} = \langle \mathbf{R}, \mathbf{L}, \mathbf{U}, \mathbf{D}, \mathbf{F}, \mathbf{B} \rangle$. As described earlier, $\mathcal{R}$ is clearly isomorphic to a subgroup of S_{48} due to the centers of each face being unaffected by the generators. Finally, note that this group perspective blurs the difference between sequences of moves and states of the cube; every move corresponds to a state of the cube (the result of applying the move to the solved state), and every state reachable from the solved state by moves corresponds to an equivalence class of all move sequences reaching that state.

► **Definition 3** (Cayley Graph). *Given a group $G = (A, \star)$ and a set of generators $S \subseteq A$, the Cayley graph G_S is the graph whose vertex set is A and where two vertices $u, v \in G$ are adjacent if there exists a generator $s \in S$ such that $u \star s = v$.*

We will use $G_{\mathcal{R}}$ to denote the Cayley graph of the Rubik's cube group under the following set of generators:

$$S_{\mathcal{R}} := \{\mathbf{R}, \mathbf{R}', \mathbf{R}2, \mathbf{L}, \mathbf{L}', \mathbf{L}2, \mathbf{U}, \mathbf{U}', \mathbf{U}2, \mathbf{D}, \mathbf{D}', \mathbf{D}2, \mathbf{F}, \mathbf{F}', \mathbf{F}2, \mathbf{B}, \mathbf{B}', \mathbf{B}2\}.$$

Note that the reason we include e.g., $\mathbf{R}'$ and $\mathbf{R}2$ in $S_{\mathcal{R}}$ is that we want to count those as single moves of the cube. Counting $\mathbf{R}2$ as 2 moves leads to a different metric, usually called the *Quarter-turn Metric*, where *God's number* is 26 [18].

► **Definition 4** (Diameter). *The diameter D of a graph $G = (V, E)$ is the maximum distance between any two vertices, where the distance between two vertices $u, v \in V$ is the length of the shortest path between them. That is,*

$$D = \max_{u, v \in \binom{V}{2}} d(u, v).$$

In this language, “God's number” is simply the diameter of the Rubik's cube graph $G_{\mathcal{R}}$.

31:6 A Demigod's Number for the Rubik's Cube

► **Definition 5** (Mean distance). *The mean distance μ of a graph $G = (V, E)$ is the average distance between any two vertices, that is,*

$$\mu = \frac{1}{\binom{|V|}{2}} \sum_{u, v \in \binom{V}{2}} d(u, v).$$

► **Definition 6** (Graph automorphism). *An automorphism over a graph $G = (V, E)$ is a bijection $\varphi : V \rightarrow V$ such that for any two vertices $u, v \in V$, we have that $(u, v) \in E$ if and only if $(\varphi(u), \varphi(v)) \in E$.*

Using the previous definition repeatedly leads to the following trivial lemma.

► **Lemma 7.** *If φ is an automorphism over a graph $G = (V, E)$, then for any two vertices $u, v \in V$, we have that $d(u, v) = d(\varphi(u), \varphi(v))$.*

► **Definition 8** (Vertex Transitivity). *A graph $G = (V, E)$ is vertex-transitive if for any two vertices $u, v \in V$, there exists an automorphism φ such that $\varphi(u) = v$.*

We now state a folklore idea that will be key to our analysis of the Rubik's cube graph.

► **Lemma 9.** *Every Cayley graph is vertex-transitive, and in particular, the Rubik's cube graph $G_{\mathcal{R}}$ is vertex-transitive.*

Proof. Let $G = (A, \star)$ be a group and $S \subseteq A$ a set of generators. We must prove that there is an automorphism φ such that $\varphi(u) = v$ for any two vertices $u, v \in A$. Let us define

$$\varphi : A \rightarrow A, \quad \varphi(x) = v \star u^{-1} \star x.$$

This definition directly implies $\varphi(u) = v$, and φ is clearly bijective since $x \mapsto u \star v^{-1} \star x$ is an inverse for φ . It remains to prove that for any two vertices $x, y \in A$, we have that $(x, y) \in E$ if and only if $(\varphi(x), \varphi(y)) \in E$. Where, by definition of the Cayley graph, a pair of vertices (a, b) is in E if there exists a generator $s \in S$ such that $a \star s = b$, and equivalently, if $a^{-1} \star b \in S$. Now observe that

$$\begin{aligned} \varphi(x)^{-1} \star \varphi(y) &= (v \star u^{-1} \star x)^{-1} \star (v \star u^{-1} \star y) \\ &= (x^{-1} \star u \star v^{-1}) \star (v \star u^{-1} \star y) \\ &= (x^{-1} \star u) \star (v^{-1} \star v) \star (u^{-1} \star y) \\ &= x^{-1} \star (u^{-1} \star u) \star y = x^{-1} \star y, \end{aligned}$$

from where we conclude by noting that

$$\begin{aligned} (x, y) \in E &\iff x^{-1} \star y \in S \\ &\iff \varphi(x)^{-1} \star \varphi(y) \in S \\ &\iff (\varphi(x), \varphi(y)) \in E. \end{aligned} \quad \blacktriangleleft$$

We conclude this section with a simple lemma stating that in a vertex-transitive graph, given that all nodes are “essentially the same”, we can think of the mean distance as the average distance from a fixed node, instead of between all pairs.

► **Lemma 10.** *Let x be any vertex in a vertex-transitive graph G . Then we have $\mu = \frac{\sum_{v \in V} d(x, v)}{|V|-1}$.*

Proof. First, note that by definition of mean distance, and using $d(u, u) = 0$, we have

$$\mu = \frac{1}{\binom{|V|}{2}} \sum_{u, v \in \binom{V}{2}} d(u, v) = \frac{1}{|V|(|V| - 1)} \sum_{u \in V} \sum_{v \in V} d(u, v).$$

Because of vertex-transitivity, for any vertex $u \in V$, there exists an automorphism φ_u such that $\varphi_u(u) = x$. Therefore, using Lemma 7, we have

$$\mu = \frac{1}{|V|(|V| - 1)} \sum_{u \in V} \sum_{v \in V} d(u, v) = \frac{1}{|V|(|V| - 1)} \sum_{u \in V} \sum_{v \in V} d(x, \varphi_u(v))$$

But as $\varphi_u : V \rightarrow V$ is a bijection for every u , we have $\sum_{v \in V} d(x, \varphi_u(v)) = \sum_{v \in V} d(x, v)$, and thus

$$\mu = \frac{1}{|V|(|V| - 1)} \sum_{u \in V} \sum_{v \in V} d(x, v) = \frac{|V|}{|V|(|V| - 1)} \sum_{v \in V} d(x, v) = \frac{\sum_{v \in V} d(x, v)}{|V| - 1}. \quad \blacktriangleleft$$

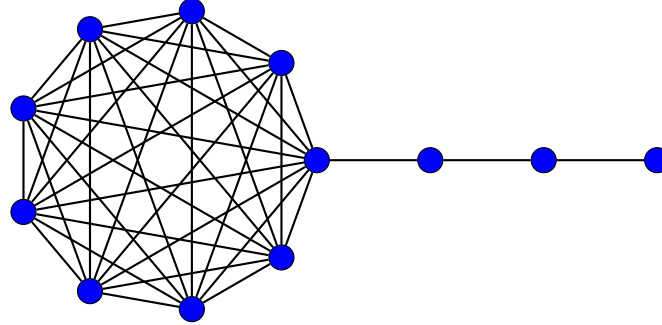
3 The Relationship Between Diameter and Mean Distance

In this section, we explore the relationship between the diameter and the mean distance of a graph, and show that for vertex-transitive graphs the diameter is at most twice the mean distance. While this result is implied by [7], we offer a more elementary exposition.

First, let us note that in arbitrary graphs, the diameter D can be much larger than the mean distance μ .

► **Proposition 11** (Folklore, cf. [22]). *For every n , there are graphs on n vertices such that $D/\mu = \Omega(n^{1/2})$.*

Proof. We can construct a graph G by taking a clique on n vertices and attaching to it a path on $n^{1/2}$ vertices, as illustrated in Figure 3.



■ **Figure 3** A graph G with $n = 9$, illustrating the proof of Proposition 11.

The diameter of this graph is $D = n^{1/2} + 1$. Noting that $|V| = n + n^{1/2} = \Omega(n)$, the mean distance can be calculated as follows:

$$\begin{aligned} \mu &= \frac{1}{\binom{|V(G)|}{2}} \sum_{u, v \in \binom{V(G)}{2}} d(u, v) \\ &= \frac{1}{\Omega(n^2)} \left(\underbrace{1 \cdot O(n^2)}_{\text{within clique vertices}} + \underbrace{O(n^{1/2}) \cdot O(n)}_{\text{within path vertices}} + \underbrace{O(n^{1/2}) \cdot n \cdot n^{1/2}}_{\text{clique-to-path}} \right) \\ &= O(1). \end{aligned} \quad \blacktriangleleft$$

31:8 A Demigod's Number for the Rubik's Cube

Furthermore, Wu et al. proved that this bound is asymptotically tight, meaning that $D/\mu = O(n^{1/2})$ [22]. It turns out, however, that such a gap between D and μ is not possible in vertex-transitive graphs, where D and μ are always a constant factor away.

► **Theorem 12.** *For any vertex-transitive graph G of diameter D and mean distance μ we have $D < 2\mu$.*

Proof. Let u, v be any pair of vertices such that $d(u, v) = D$, and use Lemma 10 to write

$$\mu = \frac{\sum_{x \in V} d(u, x)}{|V| - 1} = \frac{\sum_{x \in V} d(v, x)}{|V| - 1}, \quad (1)$$

from where

$$\begin{aligned} 2\mu &= \frac{\sum_{x \in V} d(u, x)}{|V| - 1} + \frac{\sum_{x \in V} d(v, x)}{|V| - 1} \\ &\geq \frac{\sum_{x \in V} d(u, v)}{|V| - 1} && \text{(Triangle inequality)} \\ &= \frac{|V| \cdot D}{|V| - 1} > D. \end{aligned} \quad \blacktriangleleft$$

We can see that this is tight by considering a cycle on $2n$ vertices, where $D = n$ and using Lemma 10,

$$\mu = \frac{1}{2n-1} \sum_{v \in V} d(u, v) = \frac{1}{2n-1} \left(2 \left(\sum_{i=1}^{n-1} i \right) + n \right) = \frac{2 \cdot \frac{n(n-1)}{2} + n}{2n-1} = \frac{n^2}{2n-1} = \frac{n}{2} + o(1).$$

Similarly, for the hypercube graph Q_n , we have $D = n$ and

$$\begin{aligned} \mu &= \frac{1}{2^n - 1} \sum_{v \in V} d(u, v) = \frac{1}{2^n - 1} \sum_{k=1}^n \sum_{\substack{v \in V, \\ d(u, v) = k}} k \\ &= \frac{1}{2^n - 1} \sum_{k=1}^n k \binom{n}{k} = \frac{n 2^{n-1}}{2^n - 1} = \frac{n}{2} + o(1). \end{aligned}$$

In fact, we can prove a general result that encompasses both of the previous examples: a class of graphs that is even “more symmetric” than vertex-transitive have $D \sim 2\mu$. Let us introduce the required definitions.

► **Definition 13** (Distance Transitivity). *A graph $G = (V, E)$ is distance transitive, if for every (not necessarily distinct) four vertices $u, v, x, y \in V$ with $d(u, v) = d(x, y)$, there exists an automorphism φ of G such that $\varphi(u) = x$ and $\varphi(v) = y$.*

Note that vertex-transitivity implies distance-transitivity (by taking $u = x$ and $v = y$), but the converse is not true.

► **Definition 14** (Antipodal graphs). *Given a graph $G = (V, E)$ of diameter D , we say a vertex $v \in V$ is antipodal to a vertex $u \in V$ if $d(u, v) = D$. A graph is said to be antipodal if for every vertex $u \in V$ there exists exactly one vertex $v \in V$ that is antipodal to u .*

We can now state the following result, whose proof is presented in Appendix B to avoid disrupting the flow of the main text.

► **Theorem 15.** *Let G be an antipodal distance-transitive graph of diameter D and mean distance μ . Then, $D = \lceil 2\mu - 1 \rceil$.*

4 The Demigod Number for the Rubik's Cube

Theorem 12 allows us to translate upper bounds for the average distance into upper bounds for the diameter. This is particularly useful, as the average distance is easier to certify with high confidence than the diameter. In order to provide an upper bound of the average distance we will follow the following strategy:

- We will sample a large number (500 000) of uniformly random states of the Rubik's Cube.
- For each state, we use an efficient solver³ to obtain an upper bound on the distance, which is certified by the move sequence that the solver outputs.
- We use a simple concentration bound to argue that the empirical average of the distances is a good estimate of the true average distance.

Indeed, let us prove Theorem 1 right away, after which we will explain the precise methodology used to efficiently obtain a diameter upper bound with high confidence. Besides Theorem 17, we need the following standard concentration inequality.

► **Lemma 16** (Hoeffding's inequality). *Let $X_1, X_2, \dots, X_s$ be independent random variables, all with expectation μ , and such that $X_i \in [0, C]$. Then, if $\hat{\mu} = \frac{1}{s} \sum_{i=1}^s X_i$, we have for every $t > 0$ that*

$$\Pr[\hat{\mu} \geq \mu + t] \leq \exp\left(\frac{-2st^2}{C^2}\right).$$

► **Theorem 1.** *Given a state s of the Rubik's Cube, let $d(s)$ be the distance from s to the solved state. Let S be a set of states of the Rubik's Cube sampled uniformly at random, and let $\hat{\mu}_S = \frac{1}{|S|} \sum_{s \in S} d(s)$ be the random variable corresponding to the average distance between states in S and the solved state. Then, if D denotes the diameter of the Rubik's Cube, we have*

$$\Pr_S[D \geq 2\hat{\mu}_S + 0.36] < \exp\left(\frac{-|S|}{648\,534}\right).$$

Proof. Let $G_{\mathcal{R}}$ be the Rubik's Cube graph, which is vertex-transitive by Lemma 9. Let s^* denote the vertex of $G_{\mathcal{R}}$ corresponding to the solved state. Then, let $X_1, \dots, X_{|S|}$ be i.i.d random variables corresponding to the distance between a uniformly random vertex of $G_{\mathcal{R}}$ and s^* . Using Lemma 10 we have $\mathbb{E}[X_i] = \mu$, and by Lemma 2, we have that $X_i \in [0, 205]$. Also, $\hat{\mu}_S = (\sum_{i=1}^{|S|} X_i)/|S|$. Therefore, applying Lemma 16 with $t = 0.18$, we get

$$\Pr[\mu \geq \hat{\mu}_S + 0.18] \leq \exp\left(\frac{-2|S| \cdot 0.18^2}{205^2}\right) < \exp\left(\frac{-|S|}{648\,534}\right).$$

As $D < 2\mu$ by Theorem 12, we conclude

$$\Pr[D \geq 2\hat{\mu}_S + 0.36] \leq \Pr[2\mu \geq 2\hat{\mu}_S + 0.36] < \exp\left(\frac{-|S|}{648\,534}\right). \quad \blacktriangleleft$$

³ We use <https://github.com/efrantar/rob-twophase> since it was the fastest solver we could find online.

4.1 Randomly sampling cubes

To obtain our estimates on μ_{close} and to determine we need an effective procedure for sampling states of the Rubik's Cube uniformly at random. While in a general Cayley graph this is not necessarily efficient, with the study of mixing times of random walks being a hairy problem, the particular structure of the Rubik's Cube allows us to sample uniformly at random in a very efficient way.

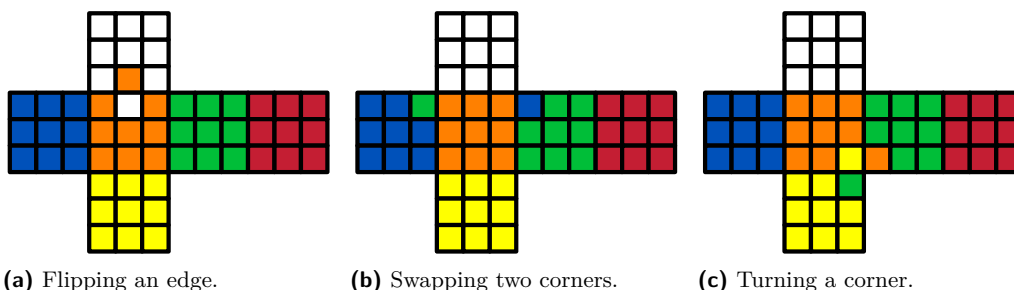
Indeed, imagine that we take out all 20 (8 corner pieces, 12 edge pieces) non-center “cubies”⁴ of the Rubik's Cube and rearrange them into a new cube. The following theorem states under what conditions such a rearrangement is a valid state of the cube (i.e., a state that can be reached from the solved state through legal moves).

► **Theorem 17** (Fundamental Theorem of Cubology, [14, Theorem 20.2.1], [15]). *A rearrangement of the cubies is valid if and only if*

- *The permutation of the corners has the same parity as the permutation of the edges. (See [14, Theorem 20.2.1] for a formal definition)*
- *The number of corners that are oriented clockwise equals the number of corners that are oriented counterclockwise modulo three. (See Figure 4c for an example of corner orientation.)*
- *The number of flipped edges is even. (See Figure 4a for an example of flipping an edge.)*

Interestingly, Theorem 17 implies an efficient algorithm for sampling. Consider the following three non-valid operations in the cube:

1. Flip the edge between **F** and **U** (see Figure 4a).
2. Swap the corners at the intersection of **F**, **U**, **L** and **F**, **U**, **R** (see Figure 4b).
3. Turn clockwise the corner at the intersection of **D**, **U**, **R** (see Figure 4c).



■ **Figure 4** Three invalid operations.

Performing these three operations leads to 12 cube configurations (there are 2 choices for flipping the edge, 2 choices for swapping the corners, and 3 choices for turning the corner). Theorem 17 implies that out of these 12 configurations, exactly one of them is valid, *no matter the state the rest of the cube is in*. Thus, the algorithm that reassembles the cube at random and rejects invalid states by Theorem 17, takes (in expectation), 12 reassembles to uniformly sample a state of the cube. Furthermore, the algorithm that reassembles the cube at random and fixes an invalid state by performing the three operations of Figure 4 takes only 1 re-assembly of the cube to sample uniformly.

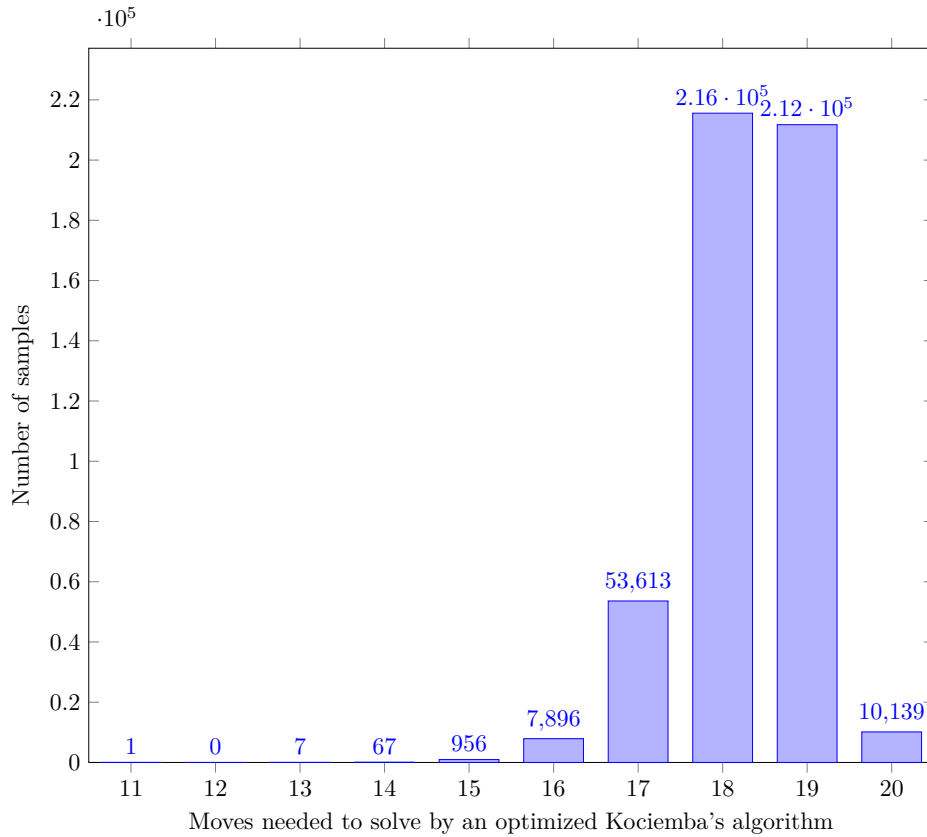
⁴ A *cubie* is a rather standard term to denote one of the pieces of physical Rubik's cube, which involves two stickers if it is an edge, and three if it is a corner.

In summary, while disassembling a Rubik’s Cube and reassembling into the solved state is considered a rather lame way of solving the cube, reassembling it randomly (and fixing the state if it is invalid) is an efficient way to sample uniformly at random from the cube.

Finally, note that, by vertex-transitivity, sampling a random pair in the cube has the same distance distribution as sampling one state of the cube and giving the distance to a *fixed* state. We will use the solved state as the fixed state in our computations and experiments.

4.2 Experimental results

Using the aforementioned methodology, we sampled 500 000 uniformly random states of the Rubik’s Cube. Out of these, no state required more than 20 moves to be solved, and the empirical mean distance obtained was 18.3189. The process took under 5 hours on a personal computer (MacBook Pro M3, 36 GB of RAM, 16 cores). A histogram is displayed in Figure 5. Our experiments can be found in <https://github.com/bsubercaseaux/RubikDemiGodNumber/>.



■ **Figure 5** Histogram showing 500 000 samples. This plot aggregates how many moves the samples needed to be solved. No more than 20 moves were needed, and the empirical mean was 18.3189.

4.3 Obtaining the Demigod’s number

While Theorem 1 already provides us with a way of obtaining an upper bound for the diameter, its sample complexity is not very practical. To obtain a probability of error of, say, 0.01%, we could set

$$\exp(-|S|/648\,534) \leq 0.0001 \implies |S| \geq 648\,534 \log(10000) \approx 6 \cdot 10^6,$$

31:12 A Demigod's Number for the Rubik's Cube

from where running a solver on each sample for 0.2 seconds, on 8 parallel threads, would take a couple of days. Getting to a probability of error of 10^{-10} would require about a year. Therefore, we now apply a simple method to reduce the sample complexity. We begin by showing that it is highly unlikely that many states of the cube are “far apart”. To this end, we say that a pair of states of the cube is “far apart” if their distance is 21 or more, otherwise, we say that the states are “close”. We then consider the following statement:

$$\phi := \text{“at least 0.03\% of the pairs of states of the cube are far apart.”} \quad (2)$$

If ϕ is true, then by sampling enough pairs of states at random we would expect to observe a pair of states that are far apart. We formalize this idea in the following lemma whose proof is straightforward.

► **Lemma 18.** *If the statement ϕ (from (2)) is true, then the probability of sampling s pairs of cube states uniformly at random and observing 0 pairs of states that are far apart is at most $(1 - 0.0003)^s$.*

We now consider the mean distance between close vertices, μ_{close} ; that is,

$$\mu_{\text{close}} := \sum_{\substack{u, v \in \binom{V}{2} \\ u, v \text{ close}}} d(u, v).$$

Moreover, we consider the empirical mean between close vertices, which we denote by $\widehat{\mu_{\text{close}}}$; that is, we sample pairs at random, compute the distance, discard the results whenever the pairs are far apart, and average the results. Then, we can use Lemma 16 with $C = 20$ to state that μ_{close} and $\widehat{\mu_{\text{close}}}$ are close to each other with high probability, obtaining the following lemma.

► **Lemma 19.** *Let $\widehat{\mu_{\text{close}}}(s)$ be the empirical mean distance over s samples. Then,*

$$\Pr [\widehat{\mu_{\text{close}}}(s) \geq \mu_{\text{close}} + 0.1] \leq \exp(-0.00005 \cdot s).$$

By Lemma 18, if we assume statement ϕ , then the probability of observing no pair of states that are far apart when sampling 500 000 pairs of states is at most

$$(1 - 0.0003)^{500\,000} < 7.02 \cdot 10^{-66}.$$

However, that is exactly what we observed, and it can be easily certified by exhibiting the sequences of moves that transform one state into the other (included in our supplementary material). Therefore, we have very significant empirical evidence (i.e., $\geq 1 - 10^{-65}$ confidence interval) for the statement ϕ being false. Equivalently, we have empirical evidence for:

$$\text{“at most 0.03\% of the pairs of states of the cube are far apart.”} \quad (3)$$

Furthermore, the empirical mean observed with 500 000 samples was 18.3189.

Since $\exp(-0.00005 \cdot 500\,000) \approx 1.38 \cdot 10^{-11}$, by Lemma 19, we have *significant empirical evidence* that

$$\mu_{\text{close}} \leq 18.3189 + 0.1 = 18.4189. \quad (4)$$

Combining facts (3) and (4) with Lemma 2, we have significant empirical evidence that

$$\mu \leq \mu_{\text{close}} + 0.03\% \cdot 205 \leq 18.4189 + 0.0615 = 18.4804. \quad (5)$$

As $2 \cdot 18.4804 = 36.9608$, using a union bound, and Theorem 12, the probability of seeing our empirical observations if the Rubik's Cube diameter were to be 37 or larger, is no more than

$$1.38 \cdot 10^{-11} + 7.02 \cdot 10^{-66} < 10^{-10},$$

and thus we have shown how to obtain a high degree of confidence with relatively few samples. Following this same approach for 90 minutes of computation results in a probability of error of roughly 10^{-4} , and thus during a lecture one could convince an audience of a 1-in-10 000 chance that the diameter of the Rubik's Cube is larger than 36.

5 Discussion

We have presented a novel approach to bounding the number of moves required to solve any state of the $3 \times 3 \times 3$ Rubik's Cube, which relies on three simple aspects of the cube: (i) its vertex-transitivity, (ii) our ability to efficiently sample uniformly random states from it, and (iii) good empirical algorithms. Because these properties extend to a variety of combinatorial puzzles (both to puzzles in the Rubik's family, such as the *Piraminx* or the *Megaminx*, as well as unrelated puzzles like the 15-puzzle on a torus), we expect that this methodology can be used to obtain similar results for other puzzles without spending the computational resources required to fully explore their state graphs. Moreover, while God's number is known for the $3 \times 3 \times 3$ cube, it remains widely open for larger puzzles [20] (with a gap larger than by a factor of 2 between lower bound and upper bound already for the $5 \times 5 \times 5$ cube), where our approach could be useful provided a good algorithm without requiring theoretical guarantees on it. In terms of related work, So Hirata has recently released two interesting papers concerning the diameter of Rubik's puzzles [8, 9], which attack the problem from different angles; either considering the girth of the cube's graph or using estimations for its branching factor. On a more theoretical line of work, Demaine et al. proved that computing the diameter of an $N \times N \times N$ Rubik's cube is NP-hard [6]⁵, and that the diameter of the $N \times N \times N$ cube is $\Theta\left(\frac{N^2}{\log N}\right)$ [5].

A particular characteristic of our approach is that our result has an intermediate epistemic status between a theorem and a heuristic. The situation, more in general, is closely related to the question of how much power randomness gives to computation. In the context of mathematical results, such a question may be phrased as follows:

Are there properties of finite mathematical objects that can only be certified efficiently to a high degree of confidence by probabilistic algorithms, but that we never can be certain of through a short proof?

For instance, consider primality testing; whether a given number is prime or not is a fully deterministic fact, in the same way as the average distance of the Rubik's Cube graph is. However, in order to practically obtain knowledge of such deterministic facts, we leverage the computational benefit of randomness, which allows us (at least in current practice), to determine facts that otherwise would be out of reach. The cost, however, is the possibility of error in the associated randomized algorithms, which forbids us from claiming to have definite proofs of the facts of interest. As usual, we can get such probability of error to be as

⁵ More in general, computing the diameter of a Cayley graph is NP-hard given a group presentation [3].

small as we deem necessary for convincing ourselves, at a modest computational price, while keeping the curse of never reaching 100% confidence. An interesting (albeit brief) discussion of this general issue can be found in Gil Kalai's blog [12].

Going back to our case, we have shown that, if one assumes momentarily that the diameter of the Rubik's Cube graph is larger than 36, then observing an empirical mean distance of around 18.3 over 500 000 samples, none of which required more than 20 moves, has probability under 10^{-10} . We encourage the readers to reproduce this computation by themselves, which should take less than a day in any modern computer. We believe this makes an extremely compelling case for the diameter of the Rubik's Cube graph being at most 36 while using a fraction of the computation required by previous approaches. Moreover, we hope that this same line of attack can be useful for analyzing other puzzles or graphs.

References

- 1 Christoph Bandelow. *Inside Rubik's Cube and Beyond*. Birkhäuser Boston, 1982. doi:10.1007/978-1-4684-7779-5.
- 2 Jonathan Borwein and David Bailey. *Mathematics by Experiment: Plausible Reasoning in the 21st Century*. A K Peters/CRC Press, New York, 2 edition, 2011. doi:10.1201/b10704.
- 3 Jin-Yi Cai, George Havas, Bernard Mans, Ajay Nerurkar, Jean-Pierre Seifert, and Igor Shparlinski. On Routing in Circulant Graphs. In Takano Asano, Hideki Imai, D. T. Lee, Shin-ichi Nakano, and Takeshi Tokuyama, editors, *Computing and Combinatorics*, pages 360–369, Berlin, Heidelberg, 1999. Springer. doi:10.1007/3-540-48686-0_36.
- 4 JJ Chen. Group theory and the Rubik's cube, 2004.
- 5 Erik D. Demaine, Martin L. Demaine, Sarah Eisenstat, Anna Lubiw, and Andrew Winslow. Algorithms for solving rubik's cubes. In Camil Demetrescu and Magnús M. Halldórsson, editors, *Algorithms - ESA 2011 - 19th Annual European Symposium, Saarbrücken, Germany, September 5-9, 2011. Proceedings*, volume 6942 of *Lecture Notes in Computer Science*, pages 689–700. Springer, 2011. doi:10.1007/978-3-642-23719-5_58.
- 6 Erik D. Demaine, Sarah Eisenstat, and Mikhail Rudoy. Solving the Rubik's cube optimally is np-complete. In Rolf Niedermeier and Brigitte Vallée, editors, *35th Symposium on Theoretical Aspects of Computer Science, STACS 2018, February 28 to March 3, 2018, Caen, France*, volume 96 of *LIPICs*, pages 24:1–24:13. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPICs.STACS.2018.24.
- 7 Mark Herman and Jonathan Pakianathan. On the distribution of distances in homogeneous compact metric spaces. *Topology and its Applications*, 193:97–99, 2015. doi:10.1016/j.topol.2015.06.011.
- 8 So Hirata. Graph-theoretical estimates of the diameters of the Rubik's cube groups, 2024. doi:10.48550/arXiv.2407.12961.
- 9 So Hirata. Probabilistic estimates of the diameters of the Rubik's cube groups, 2024. doi:10.48550/arXiv.2404.07337.
- 10 Nathan Jacobson. *Basic Algebra I*. Dover Books on Mathematics. Dover Publications, Mineola, NY, 2 edition, June 2009.
- 11 JPerm. Learn How to Solve a Rubik's Cube in 10 Minutes (Beginner Tutorial). <https://www.youtube.com/watch?v=7Ron6MN45LY>, 2019. [Accessed 03-08-2024].
- 12 Gil Kalai. The probabilistic proof that $2^{400} - 593$ is a prime: a revolutionary new type of mathematical proof, or not a proof at all? <https://gilkalai.wordpress.com/2021/04/22/the-probabilistic-proof-that-2400-593-is-a-prime-a-revolutionary-new-type-of-mathematical-proof-or-not-a-proof-at-all/>, 2011. [Accessed 23-04-2025].
- 13 Herbert Kociemba. Close to God's algorithm. In *Cubism for Fun*, pages 10–13, 1992.
- 14 James T. Mulholland. Math 302: Combinatorial mathematics. <http://www.sfu.ca/~jtmulhol/math302/notes/302notes.pdf>. [Accessed 08-08-2024].

- 15 James T. Mulholland. Permutation puzzles - Rubik's cube. <https://www.sfu.ca/~jtmulhol/math302/puzzles-rc-cubology.html>. [Accessed 08-08-2024].
- 16 G. Polya. *Mathematics and Plausible Reasoning, Volume 1: Induction and Analogy in Mathematics*. Princeton University Press, 1954. doi:10.2307/j.ctv14164db.
- 17 Tomas Rokicki, John Dethridge, Herbert Kociemba, and Morley Davidson. God's Number is 20. <https://www.cube20.org/>. [Accessed 08-08-2024].
- 18 Tomas Rokicki, John Dethridge, Herbert Kociemba, and Morley Davidson. God's Number is 26 in the Quarter Turn Metric. <https://cube20.org/qtm/>. [Accessed 08-08-2024].
- 19 Tomas Rokicki, Herbert Kociemba, Morley Davidson, and John Dethridge. The Diameter of the Rubik's Cube Group Is Twenty. *SIAM Review*, 56(4):645–670, 2014. doi:10.1137/140973499.
- 20 Speedsolving.com Wiki. God's Algorithm - Optimal solutions of the Rubik's Cube. https://www.speedsolving.com/wiki/index.php/God%27s_Algorithm#Table_of_God.27s_Numbers, 2024. [Accessed 06-08-2024].
- 21 Douglas B. West. Problems in Graph Theory and Combinatorics. <http://dwest.web.illinois.edu/openp/>. [Accessed 02-08-2024].
- 22 Baoyindureng Wu, Guojie Liu, Xinhui An, Guiying Yan, and Xiaoping Liu. A conjecture on average distance and diameter of a graph. *Discrete Mathematics, Algorithms and Applications*, 03(03):337–342, 2011. doi:10.1142/S1793830911001255.

A The Beginner's Method and the “Human's Number”

In order to make this article self-contained, we provide a brief overview of the “*beginner's method*” to solve the Rubik's Cube⁶, from which we have the following:

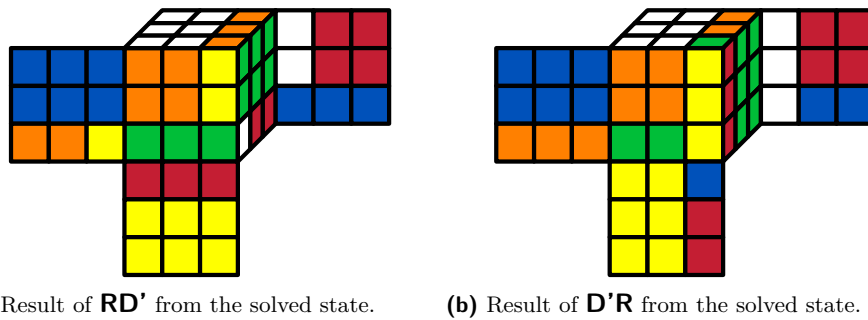
► **Lemma 2** (Human's Number). *Any position of the Rubik's Cube can be solved in at most 205 moves.*

In a nutshell, the beginner's method consists of solving the Rubik's Cube by “*layers*”, as opposed to by faces. Before we begin with its exposition, however, it is worth establishing some notation for the different Rubik's Cube moves. We will use “Singmaster” notation, credited to British mathematician David Singmaster. To specify a turn on a face, we use the first letter of the face's name: **R** for the right face, **L** for the left face, **U** for the upper face, **D** for the down face, **F** for the front face, and **B** for the back face. If the face is to be rotated by 90° clockwise, we add no suffix, e.g., **R** means a clockwise rotation by 90° of the right face. For a counterclockwise rotation by 90°, we add a prime symbol, e.g., **D'** means a counterclockwise rotation by 90° of the down face. For a 180° rotation, we add a 2 after the letter, e.g., **F2** means a 180° rotation of the front face. A proper formalization of what a “*move*” actually is can be found in Section 2, where we view the Rubik's Cube as a group. As an example to check our understanding of the notation, the non-commutativity of the Rubik's Cube group is evidenced in Figure 6.

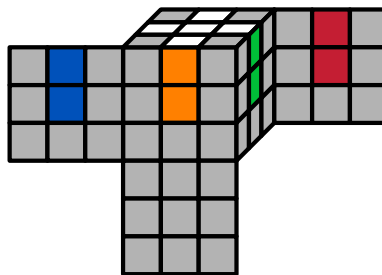
Step 1: The white cross

The first step consists of creating a “*cross*” on one face of the cube, which we will assume to be white without loss of generality. To achieve this, one must move every white “edge” (i.e., a piece with two colors) to the correct position. That is, e.g., the white-orange edge must be placed so that its white sticker is adjacent to the white center and its green sticker is adjacent to the orange center. We illustrate the result of this step in Figure 7, and a conservative

⁶ We encourage, nonetheless, the interested reader to look into the many YouTube videos (e.g., [11]) that guide the process.



■ **Figure 6** Illustration of the non-commutativity of the Rubik's Cube.

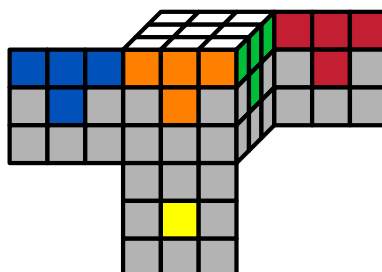


■ **Figure 7** The white cross is solved.

bound is that this can always be achieved in 20 moves, as each of the 4 white edges to place can always be placed in at most 5 moves or fewer. This step can be done intuitively, that is, without memorizing any particular algorithm⁷.

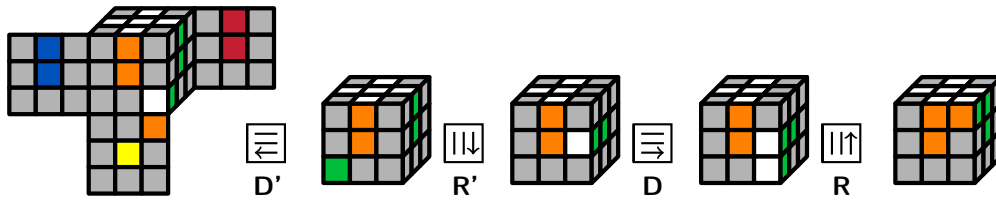
Step 2: White corners

The second step consists of placing the white corners in their correct position, one by one. To place a white corner, one can first bring it to the opposite layer (i.e., the bottom layer, whose center is yellow), and then proceed according to a handful of cases, one of which is illustrated in Figure 9. The result of this step is illustrated in Figure 8. Conservatively, this step can always be achieved in 15 moves per corner, and 60 in total. This accounts for the cases when a white corner is in the correct location but oriented incorrectly, in which case a non-white corner can be placed in that spot, thus allowing the white corner to be placed in the correct orientation afterward.

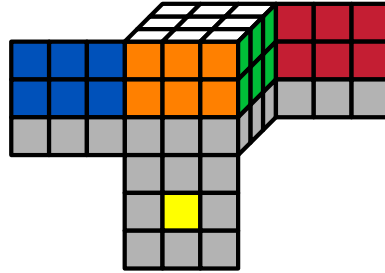


■ **Figure 8** The white corners are solved, and thus the first layer is completed.

⁷ In the Rubik's Cube literature, a move sequence with a concrete purpose (e.g., permuting 3 corner pieces) is traditionally called an "algorithm".



■ **Figure 9** Illustration of one of the cases for placing a corner in the first layer (Step 2 of the beginner's method), through the move sequence **D'R'DR**.



■ **Figure 10** The first two layers are solved.

Step 3: Edges of the second layer

The third step consists of placing the edges of the second layer in their final position, as illustrated in Figure 10. For instance, the orange-green edge must be placed so that its orange sticker is adjacent to the orange center and its green sticker is adjacent to the green center. The main algorithm to solve this step is illustrated in Figure 11. A conservative bound, again due to cases in which a misoriented edge must be first replaced before placing it in the correct orientation, is 20 moves per edge.

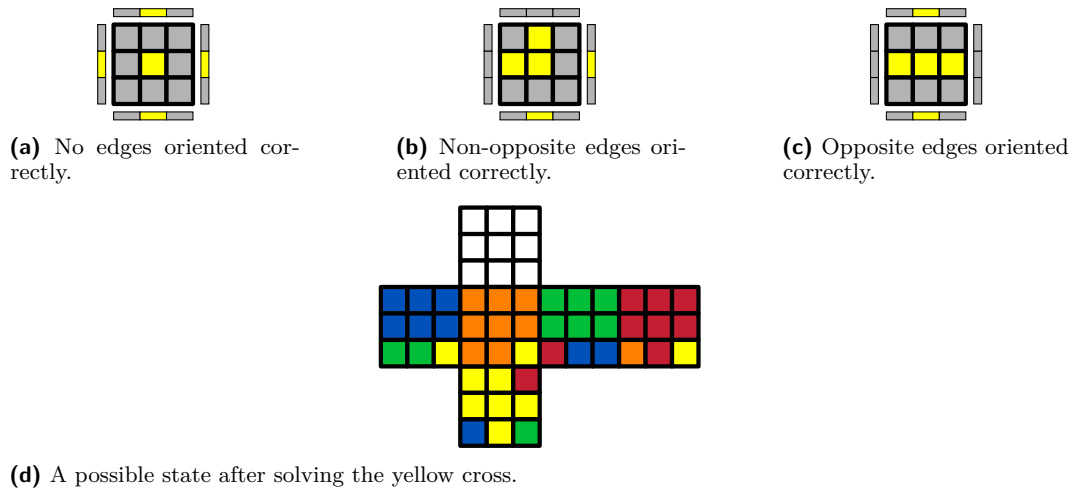
Step 4: The yellow cross

We now turn our attention to the yellow face. The goal of this step is to solve the orientation of the yellow cross, that is, to make all yellow edges have their yellow sticker adjacent to the yellow center, as depicted in Figure 12d. We may face three different scenarios in this step (if it is not already solved), as illustrated in Figure 12. These can all be solved by the same algorithm, potentially repeated according to which of the three non-solved cases we encounter. The algorithm is simply: **FRUR'U'F'**. Applying it from case 12a leads to case 12b, and applying it again leads to case 12c, from where a last application solves the yellow cross. That way, we need at most 3 applications, leading to a conservative bound of 18 moves for this step.



■ **Figure 11** Illustration of the algorithm to solve the edges of the second layer (Step 3 of the beginner's method).

31:18 A Demigod's Number for the Rubik's Cube



■ **Figure 12** Illustration of Step 4 of the beginner's method.



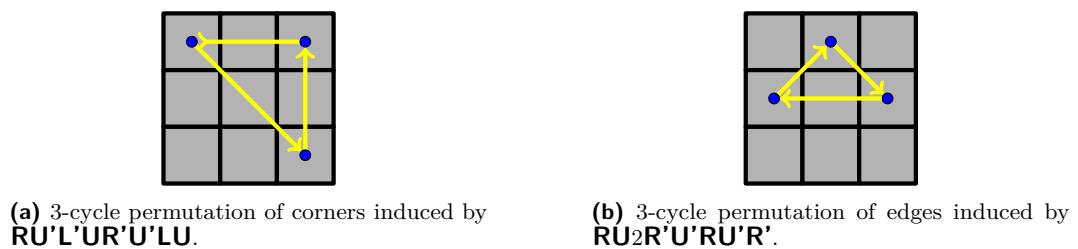
■ **Figure 13** Illustration of a case for Step 5 of the beginner's method.

Step 5: Permuting yellow edges

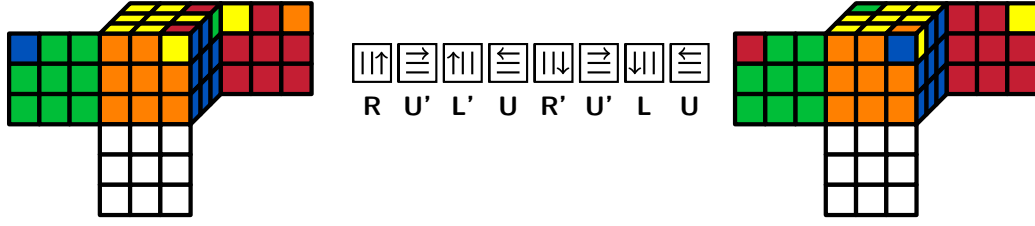
Now, we permute the yellow edges so that each of them gets to its desired position. This step can be solved by repeated application of a single algorithm, that induces a 3-cycle of the yellow edges, as illustrated in Figures 13 and 14b. As this algorithm is applied at most 3 times, we have a conservative bound of 21 moves for this step.

Step 6: Permuting yellow corners

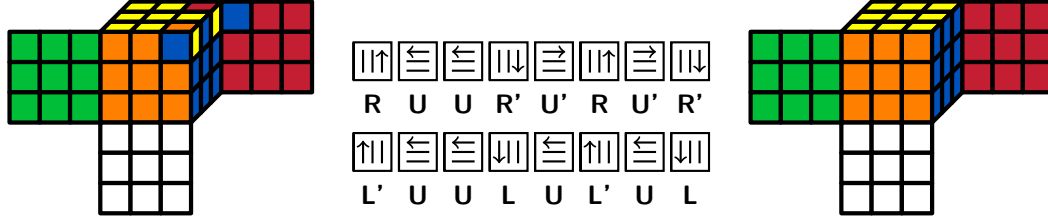
This step is analogous to the previous one but over the corners; we permute the yellow corners so that each of them gets to its desired position. This step can also be solved by repeated application of a single algorithm, that induces a 3-cycle of the yellow corners, as illustrated in Figures 14a and 15. This algorithm is applied at most 3 times, leading to a conservative bound of 24 moves for this step.



■ **Figure 14** Illustration of the 3-cycle algorithms for permuting yellow corners and edges, corresponding to Steps 5 and 6 of the beginner's method.



■ **Figure 15** Illustration of a case for Step 6 of the beginner's method.



■ **Figure 16** Illustration of a case for the final step of the beginner's method, using the move sequence: $RU_2R'U'RU'R'L'U_2LUL'UL$.

Step 7: Orienting yellow corners

The last step consists of orienting the yellow corners, which again can be achieved by repeated applications of a single algorithm that changes the orientation of two adjacent corners (illustrated in Figure 16):

$$RU_2R'U'RU'R'L'U_2LUL'UL.$$

This algorithm needs to be applied at most 3 times, leading to a conservative bound of 42 moves for this step.

The following table summarizes the “proof” of Lemma 2:

	Step	Moves
White cross		20
White corners		60
Edges of the second layer		80
Yellow cross		18
Permuting yellow edges		21
Permuting yellow corners		24
Orienting yellow corners		42
Total		205

B Proof of Theorem 15

Before the proof, we will need a couple of auxiliary observations. Given that distance-transitive graphs are vertex-transitive, we already have $D < 2\mu$ by Theorem 12. Therefore, it will suffice to prove that $2\mu - 1 \leq D$, and then use that D must be an integer.

For a graph $G = (V, E)$, and vertex $u \in V$, let $\Gamma_u(i) := \{v \in V \mid d(u, v) = i\}$ be the set of vertices at distance exactly i from u .

► **Lemma 20.** *Let $G = (V, E)$ be an antipodal distance-transitive graph, and $u \in V$ any vertex. Then, for every $i \in \{0, \dots, D\}$, we have*

$$|\Gamma_u(i)| = |\Gamma_u(D - i)|.$$

Proof. Let v be the only vertex antipodal to u . As distance transitivity implies vertex transitivity, we have by Lemma 7, that $|\Gamma_u(i)| = |\Gamma_v(i)|$ for all $i \in \{0, \dots, D\}$. Therefore, our goal is to show that $|\Gamma_v(i)| = |\Gamma_u(D - i)|$ for all $i \in \{0, \dots, D\}$. Indeed, fix any $i \in \{0, \dots, D\}$, and let $w \in V$ be an arbitrary vertex such that $d(v, w) = i$. Then, let z be an arbitrary vertex that is in some shortest path between u and v , and such that $d(z, v) = i$. Note that as z is in a shortest path between u and v , we have $d(z, u) = D - i$. Now, due to distance transitivity, there exists an automorphism φ of G such that $\varphi(v) = v$, and $\varphi(w) = z$. By Lemma 7, we have that

$$d(u, w) = d(\varphi(u), \varphi(w)) = d(\varphi(u), z),$$

and we claim that $\varphi(u) = u$. Indeed, since $D = d(u, v) = d(\varphi(u), \varphi(v)) = d(\varphi(u), v)$, we know that $\varphi(u)$ is antipodal to v , and as u is antipodal to v , we have $\varphi(u) = u$ since G is antipodal. Therefore, we have $d(u, w) = d(u, z) = D - i$, and thus $w \in \Gamma_u(D - i)$. As w was arbitrary, we have $\Gamma_v(i) \subseteq \Gamma_u(D - i)$. Because the previous argument is symmetric in u and v , and can be used with $D - i$ instead of i , we have $\Gamma_u(D - 1) \subseteq \Gamma_v(i)$, and thus $\Gamma_v(i) = \Gamma_u(D - i)$. This concludes the proof. ◀

We are now ready to prove Theorem 15.

Proof of Theorem 15. Fix any vertex $u \in V(G)$. Using Lemmas 10 and 20, we have

$$\begin{aligned} (|V(G)| - 1)\mu &= \sum_{i=0}^D i \cdot |\Gamma_u(i)| \\ &= \sum_{i=0}^{\lfloor D/2 \rfloor} i \cdot |\Gamma_u(i)| + \sum_{i=0}^{\lceil D/2 \rceil} (D - i) \cdot |\Gamma_u(D - i)| \\ &\leq \sum_{i=0}^{\lfloor D/2 \rfloor} i \cdot |\Gamma_u(i)| + (D - i) \cdot |\Gamma_u(i)| \\ &= D \sum_{i=0}^{\lfloor D/2 \rfloor} |\Gamma_u(i)|. \end{aligned}$$

But we also have

$$|V(G)| = \sum_{i=0}^D |\Gamma_u(i)| \geq 2 \sum_{i=0}^{\lfloor D/2 \rfloor} |\Gamma_u(i)|,$$

and therefore

$$\mu \leq \frac{D \cdot |V(G)|}{2(|V(G)| - 1)} = \frac{D}{2} + \frac{D}{2(|V(G)| - 1)} \leq \frac{D}{2} + \frac{1}{2},$$

from where we can easily conclude. Indeed, $\mu \leq \frac{D}{2} + \frac{1}{2}$ implies $2\mu - 1 \leq D$. Since D is an integer, this implies $\lceil 2\mu - 1 \rceil \leq D$, and by Theorem 12, we have

$$D < 2\mu \leq \lceil 2\mu \rceil \leq \lceil 2\mu - 1 \rceil + 1.$$

Hence, D is an integer such that $\lceil 2\mu - 1 \rceil \leq D < \lceil 2\mu - 1 \rceil + 1$, and thus we conclude $D = \lceil 2\mu - 1 \rceil$. ◀