Miniature 1

Fibonacci Numbers, Quickly

The **Fibonacci numbers** $F_0, F_1, F_2, \ldots$ are defined by the relations $F_0 = 0$, $F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$ for $n = 0, 1, 2, \ldots$. Obviously, F_n can be calculated using roughly n arithmetic operations.

By the following trick we can compute it faster, using only about $\log n$ arithmetic operations. We set up the 2×2 matrix

$$M := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$\begin{pmatrix} F_{n+2} \\ F_{n+1} \end{pmatrix} = M \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix},$$

and therefore,

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = M^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(we use the associativity of matrix multiplication).

For $n = 2^k$, we can compute M^n by repeated squaring, with k multiplications of 2×2 matrices. For n arbitrary, we write n in binary as $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_t}$, $k_1 < k_2 < \cdots < k_t$, and then we calculate the power M^n as $M^n = M^{2^{k_1}} M^{2^{k_2}} \cdots M^{2^{k_t}}$. This needs at most $2k_t \le 2\log_2 n$ multiplications of 2×2 matrices.

Remarks. A similar trick can be used for any sequence $(y_0, y_1, y_2, \ldots)$ defined by a recurrence $y_{n+k} = a_{k-1}y_{n+k-1} + \cdots + a_0 y_n$, where k and $a_0, a_1, \ldots, a_{k-1}$ are constants.

If we want to compute the Fibonacci numbers by this method, we have to be careful, since the F_n grow very fast. From a formula in Miniature 2 below, one can see that the number of decimal digits of F_n is of order n. Thus, we must use multiple precision arithmetic, and so the arithmetic operations will be relatively slow.

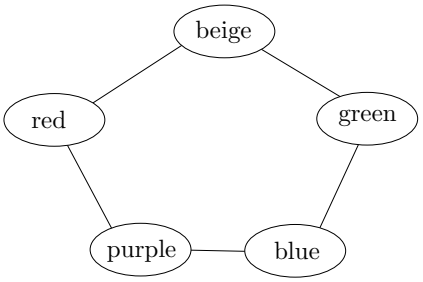
Sources. This trick is well known, but so far I haven't encountered any reference to its origin.

Miniature 28

The Secret Agent and the Umbrella

A secret government agent in a desert training camp of a terrorist group has very limited possibilities of sending messages. He has five scarves: red, beige, green, blue, and purple, and he wears one of them with his uniform every day. The analysts at the headquarters then determine the color of his scarf from a satellite photography.

But since the scarves are not really clean, it turned out that certain pairs of colors cannot be distinguished reliably. The possibilities of confusion are shown in the next picture:



For example, one cannot reliably tell purple from blue nor from red, but there is no danger of confusing purple with beige or green.

In order to transmit reliably, the agent can, for example, use only the blue and red scarves, and thereby send one of two possible messages every day—one bit in the computer science language. He can communicate one of 2^k possible messages in k days.

Among every three scarves there are some two that can be confused, and so it may seem that there is no chance to send more than one bit per day. But there is a better way! In two successive days, the agent can send one of five messages, e.g., as follows:

	the first day	the second day
message 1	red	red
message 2	beige	green
message 3	green	purple
message 4	blue	beige
message 5	purple	blue

Indeed, there is no chance of mistaking any of these two-day combinations for another, as the reader can easily check. So the agent can transmit one of $5^{k/2} = \sqrt{5}^k$ possible messages in k days (for k even), and the efficiency per day has increased from 2 to $\sqrt{5}$.

Can the efficiency be increased further using three-day or ten-day combinations, say? This is a difficult mathematical problem. The answer is no, and the following masterpiece is the only known proof.

First we formulate the problem in mathematical terms (and generalize it). We consider some **alphabet** S; in our case S consists of the five possible colors of the scarf. Some pairs of symbols of S can be confused (in other words, are *interchangeable*), and this is expressed by a graph $G = (S, E)$, where the interchangeable pairs of symbols of S are connected by edges. For the situation with five scarves, the graph is drawn in the picture on the preceding page, and it is a cycle of length 5, i.e., C_5.

Let us consider two messages of length k: a message $a_1 a_2 \cdots a_k$ and a message $b_1 b_2 \cdots b_k$. In the terminology of coding theory, these are the words of length k over the alphabet S; see Miniature 5. These messages are interchangeable if and only if a_i is interchangeable with b_i (meaning that $a_i = b_i$ or $\{a_i, b_i\} \in E$) for every $i = 1, 2, \ldots, k$.

Let $\alpha_k(G)$ be the maximum size of a set of messages of length k with no interchangeable pair. In particular, $\alpha_1(G)$ is the maximum size of an **independent set** in G, i.e., a subset of vertices in which no pair is connected by an edge. This quantity is usually denoted by $\alpha(G)$. For our example, we have $\alpha_1(C_5) = \alpha(C_5) = 2$. Our table proves that $\alpha_2(C_5) \geq 5$, and actually equality holds—the inequality $\alpha_2(C_5) \leq 5$ is a very special case of the result we are about to prove.

The **Shannon capacity** of a graph G is defined as follows:

$$\Theta(G) := \sup\left\{\alpha_k(G)^{1/k} : k = 1, 2, \ldots\right\}.$$

It represents the maximum possible efficiency of message transmission per symbol. For a sufficiently large k, the agent can send one from approximately $\Theta(C_5)^k$ possible messages in k days, and not more. We prove the following:

Theorem. $\Theta(C_5) = \sqrt{5}$.

First we observe that $\alpha_k(G)$ can be expressed as the maximum size of an independent set of a suitable graph. The vertex set of this graph is S^k, meaning that the vertices are all possible messages (words) of the length k, and two vertices $a_1a_2\cdots a_k$ and $b_1b_2\cdots b_k$ are connected by an edge if they are interchangeable. We denote this graph by G^k, and we call it the **strong product** of k copies of G.

The strong product $H \cdot H'$ of two arbitrary graphs H and H' is defined as follows:

$$\begin{aligned} V(H \cdot H') &= V(H) \times V(H'), \\ E(H \cdot H') &= \{\{(u,u'),(v,v')\} : (u = v \text{ or } \{u,v\} \in E(H)) \\ &\qquad \text{and at the same time} \\ &\qquad (u' = v' \text{ or } \{u',v'\} \in E(H'))\}. \end{aligned}$$

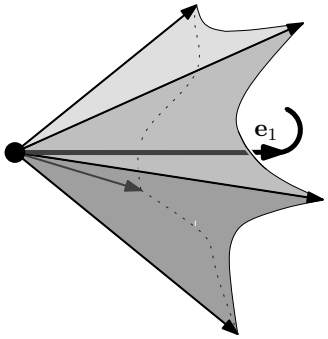
For bounding $\Theta(C_5)$, we thus need to bound above the maximum size of an independent set in each of the graphs C_5^k.

We are going to establish two general results relating independent sets in graphs to certain systems of vectors. Let $H = (V, E)$ be an arbitrary graph. An **orthogonal representation** of H is a mapping $\rho\colon V \to \mathbb{R}^n$, for some n, that assigns a *unit* vector $\rho(v)$ to every vertex $v \in V(H)$ (i.e., $\|\rho(v)\| = 1$), such that the following holds:

> If two distinct vertices u, v are *not connected* by an edge, then the corresponding vectors are *orthogonal*. In symbols, $\{u, v\} \notin E$ implies $\langle \rho(u), \rho(v) \rangle = 0$.

(We use $\langle ., . \rangle$ for the standard scalar product in $\mathbb{R}^n$.)

To prove our main theorem, we will need an interesting orthogonal representation ρ_{LU} of the graph C_5 in $\mathbb{R}^3$, the "Lovász umbrella". Let us imagine a folded umbrella with five ribs, of unit length, whose tube is the vector $\mathbf{e}_1 = (1, 0, 0)$. Now we slowly open the umbrella until all pairs of nonneighboring ribs become orthogonal:



At this moment, the ribs define unit vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_5$. By assigning the vector $\mathbf{v}_i$ to the ith vertex of the graph C_5, we get an orthogonal representation ρ_{LU}. A simple calculation yields the opening angle of the umbrella: we obtain $\langle \mathbf{v}_i, \mathbf{e}_1 \rangle = 5^{-1/4}$, which we will soon need.

Every orthogonal representation of a graph G provides an upper bound on $\alpha(G)$:

Lemma A. *If H is a graph and ρ is an orthogonal representation of H, then $\alpha(H) \le \vartheta(H, \rho)$, where*

$$\vartheta(H, \rho) := \max_{v \in V(H)} \frac{1}{\langle \rho(v), \mathbf{e}_1 \rangle^2}.$$

Proof. Producing an orthogonal representation ρ with $\vartheta(H, \rho)$ minimum has the following geometric meaning: We want to pack all the

unit vectors $\rho(\mathbf{v})$ into a spherical cap centered at $\mathbf{e}_1$ and with the smallest possible radius.

The vectors resist such a packing since pairs corresponding to nonedges must be orthogonal. In particular, the vectors corresponding to an independent set in H form an orthonormal system, and for such a system the minimum cap radius can be calculated exactly.

For a formal proof we need to know that for an arbitrary orthonormal system of vectors $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ in some $\mathbb{R}^n$ and an arbitrary vector $\mathbf{u}$, we have

$$\sum_{i=1}^{m} \langle \mathbf{v}_i, \mathbf{u} \rangle^2 \le \|\mathbf{u}\|^2.$$

Indeed, the given system $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ can be extended to an orthonormal basis $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ of $\mathbb{R}^n$, by adding $n-m$ other suitable vectors $(\mathbf{v}_{m+1}, \mathbf{v}_{m+2}, \ldots, \mathbf{v}_n)$. The ith coordinate of $\mathbf{u}$ with respect to this basis is $\langle \mathbf{v}_i, \mathbf{u} \rangle$, and we have $\|\mathbf{u}\|^2 = \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle^2$ by the Pythagorean theorem. The required inequality is obtained by omitting the last $n - m$ terms on the right-hand side.

Now if $I \subseteq V(H)$ is an independent set in H, then, as noted above, the vectors $\rho(v)$ with $v \in I$ form an orthonormal system, and so

$$\sum_{v \in I} \langle \rho(v), \mathbf{e}_1 \rangle^2 \le \|\mathbf{e}_1\|^2 = 1.$$

Hence there exists $v \in I$ with $\langle \rho(v), \mathbf{e}_1 \rangle^2 \le \frac{1}{|I|}$, and thus $\vartheta(H, \rho) \ge |I|$. □

The lemma together with the Lovász umbrella gives

$$\alpha(C_5) \le \vartheta(C_5, \rho_{\mathrm{LU}}) = \sqrt{5}.$$

This is not (yet) an earth-shattering result, since everyone knows that $\alpha(C_5) = 2$. We need to complement this with the following lemma, showing that orthogonal representations behave well with respect to the strong product.

Lemma B. *Let H_1, H_2 be graphs, and let ρ_i be an orthogonal representation of H_i, $i = 1, 2$. Then there is an orthogonal representation ρ of the strong product $H_1 \cdot H_2$ such that $\vartheta(H_1 \cdot H_2, \rho) = \vartheta(H_1, \rho_1) \cdot \vartheta(H_2, \rho_2)$.*

Applying Lemma B inductively to the strong product of k copies of C_5, we obtain

$$\alpha(C_5^k) \le \vartheta(C_5, \rho_{\mathrm{LU}})^k = \sqrt{5}^k,$$

which proves that $\Theta(C_5) \le \sqrt{5}$ and thus yields the theorem.

Proof of Lemma B. We recall the operation of the **tensor product**, already used in Miniature 18. The tensor product of two vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$ is a vector in $\mathbb{R}^{mn}$, denoted by $\mathbf{x} \otimes \mathbf{y}$, with coordinates corresponding to all products $x_i y_j$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. For example, for $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2)$, we have

$$\mathbf{x} \otimes \mathbf{y} = (x_1y_1, x_2y_1, x_3y_1, x_1y_2, x_2y_2, x_3y_2) \in \mathbb{R}^6.$$

We need the following fact, whose routine proof is left to the reader:

$$\langle \mathbf{x} \otimes \mathbf{y}, \mathbf{x}' \otimes \mathbf{y}' \rangle = \langle \mathbf{x}, \mathbf{x}' \rangle \cdot \langle \mathbf{y}, \mathbf{y}' \rangle \tag{22}$$

for arbitrary $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^m$, $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^n$.

Now we can define an orthogonal representation ρ of the strong product $H_1 \cdot H_2$ as in the lemma. The vertices of $H_1 \cdot H_2$ are pairs (v_1, v_2), $v_1 \in H_1$, $v_2 \in H_2$. We put

$$\rho((v_1, v_2)) := \rho_1(v_1) \otimes \rho_2(v_2).$$

Using (22) we can easily verify that ρ is an orthogonal representation of $H_1 \cdot H_2$, and the equality $\vartheta(H_1 \cdot H_2, \rho) = \vartheta(H_1, \rho_1) \cdot \vartheta(H_2, \rho_2)$ follows as well. This completes the proof of Lemma B. □

Remarks. The quantity

$$\vartheta(G) = \inf\{\vartheta(G, \rho) : \rho \text{ an orthogonal representation of } G\}$$

is called the **Lovász theta function** of the graph G. As we have seen, it gives an upper bound for $\alpha(G)$, the independence number of the graph. It is not hard to prove that it also provides a lower bound on the **chromatic number** of the complement of the graph G, or in other words, the minimum number of complete subgraphs needed to cover G. Computing the independence number or the chromatic number of a given graph is algorithmically hard (NP-complete), but

surprisingly, $\vartheta(G)$ can be computed in polynomial time (more precisely, approximated with arbitrary required precision). Because of this and several other remarkable properties, the theta function is very important.

The Shannon capacity of a graph is a much harder nut to crack. *No algorithm at all*, polynomial or not, is known for computing or approximating it. And we need not go far for an unsolved case—$\Theta(C_7)$ is not known! If the agent had seven scarves, nobody can tell him the best way of transmitting.

Source. L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Th. **IT-25** (1979), 1–7.