Automatic CounTilings

Doron ZEILBERGER[1]

I Have a Dream

One day it would be possible to write in English, or in an English-like super-high-level programming language, the following command:

`Write a Maple program that inputs an arbitrary positive integer k, a symbol n, and an arbitrary set` T `of tiles, as well as a symbol t, and outputs the rational function of t whose Maclaurin series's coefficient of` t^n `is the number of different tilings of a k by n rectangle by the tiles of` T.

But Yesterday's Dream Came To Pass

This day hasn't arrived yet, so I had to spend a few weeks writing such a program myself. This is the human-generated Maple package `TILINGS` accompanying this article.

But, this is already a great step forward. Once I wrote the program, my beloved servant, Shalosh B. Ekhad, can solve (potentially) infinitely many enumeration problems, completely rigorously, for enumerating tilings of rectangles of *arbitrary* width and using an *arbitrary* set of tiles (the tiles even do not have to be connected).

Traditionally, a human would have had to tackle each specific width (k) and each specific set of tiles (T), (even the set consisting of the two dimers, the vertical and horizontal 1 by 2 rectangles) *one at a time*. He or she would have had to figure out the combinatorics of the situation, establish a "structure theorem", then deduce from this a *set of equations*, and finally, solve this set of equations (this very last part, starting in the seventies, was possibly aided by Maple or Mathematica). Of course, the human will only be able to do it for very small k, and very simple sets of tiles, since very soon the structure theorem, and consequently the set of equations, would be too hard to *derive* by hand, let alone solve.

Once the meta-program above would be realized, I would become superfluous. But even at the present level, the human-made Maple program `TILINGS` can generate, in principle, infinitely many articles, and Ph.D. theses, proving general, rigorous and *interesting* results. Except for very simple cases, of course, the complexity of the proofs and results are beyond mere humans.

The Simplest Not-Completely-Trivial Example

Let $a(n)$ be the number of ways of tiling a $2 \times n$ rectangle using the horizontal and vertical dominoes,

---

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `zeilberg at math dot rutgers dot edu` ,
`http://www.math.rutgers.edu/~zeilberg` . First version: Jan. 20, 2006. Accompanied by Maple packages `TILINGS` and `RecTILINGS` downloadable from Zeilberger's website. Supported in part by the NSF.

i.e. the set of tiles is:

$$XX \quad , \quad \begin{matrix} X \\ X \end{matrix} \quad .$$

Let $\mathcal{A}(n)$ be the set of tilings of a $2 \times n$ rectangular board by the above two tiles. Of course, if $n = 0$, then $\mathcal{A}(n)$ consists of one tiling only: the **empty tiling**, but if $n > 0$ then there are two cases to consider, regarding the bottom-left cell.

Case I: It participates in the vertical tile, in which case the board looks like

$$\begin{matrix} 10000\ldots 0 \\ 10000\ldots 0 \end{matrix} \quad ,$$

where 1 means an occupied cell and 0 means a still-untiled cell.

Case II: It participates in the horizontal tile, in which case the board looks like

$$\begin{matrix} 00000\ldots 0 \\ 11000\ldots 0 \end{matrix} \quad .$$

Now the first case is clearly "isomorphic" to $\mathcal{A}(n-1)$, and the isomorphism is obtained by "chop the 1's" (i.e. remove the vertical tile), reducing the problem to that of tiling a $2 \times (n-1)$ board. But Case II requires us to introduce a new **auxiliary** set, let's call it $\mathcal{B}(n)$, that of tiling a "jagged" rectangle with the two leftmost cells of the bottom row removed.

So now we forget about the original problem, and try to find a "structure theorem" for $\mathcal{B}(n)$. Of course, if $n = 0, 1$, then $\mathcal{B}(n)$ is the empty set, but if $n \geq 2$ then, we consider the leftmost unoccupied cell of the top row. Now there is only one case to consider: it is tiled by a horizontal tile, getting a board of the form

$$\begin{matrix} 11000\ldots 0 \\ 11000\ldots 0 \end{matrix} \quad ,$$

which is trivially "isomorphic" to $\mathcal{A}(n-2)$.

We now have two "structure theorems":

$$\mathcal{A}(n) \equiv \mathcal{A}(n-1) \cup B(n) \quad (n > 0) \quad ,$$

$$\mathcal{B}(n) \equiv A(n-2) \quad . \qquad (MarkovianScheme)$$

Now taking cardinalities, calling $a(n) := |\mathcal{A}(n)|$, $b(n) := |\mathcal{B}(n)|$, we have the system of **linear recurrences**

$$a(n) = a(n-1) + b(n) \quad (n > 0) \quad ,$$

$$b(n) = a(n-2) \quad ,$$

with the obvious initial conditions $a(-2) = 0, a(-1) = 0, a(0) = 1$.

Doing the usual generatingfunctionology, we get the generating functions

$$f(t) := \sum_{n=0}^{\infty} a(n)t^n = \frac{1}{1-t-t^2} \quad ,$$

$$g(t) := \sum_{n=0}^{\infty} b(n)t^n = \frac{t^2}{1-t-t^2} \quad .$$

Equivalently, we can use **Polya's Picture Writing**[P] and use **weight-enumerators** to deduce the system

$$f(t) = 1 + tf(t) + g(t) \quad ,$$

$$g(t) = t^2 f(t) \quad ,$$

and solve the system of **two linear equations** in the **two unknowns** $f(t), g(t)$. But we can do better still: keep track of the number of occurrences of each tile. Introducing the variables h and v for a horizontal and vertical tile respectively, and defining the **weight** of a tiling to be $h^{\#horizontal\_tiles} v^{\#vertical\_tiles} t^n$, where n is its length, and defining $F(t,h,v)$ to be the sum of all the weights of all tilings, and analogously $G(t,h,v)$ for the set $\mathcal{B}$, we get the system

$$F(t,h,v) = 1 + tvF(t,H,V) + hG(t,H,V) \quad ,$$

$$G(t,h,v) = t^2 hF(t,T,V) \quad ,$$

and solving this system gives:

$$F(t,h,v) := \frac{1}{1-vt-h^2t^2} \quad ,$$

$$G(t,h,v) := \frac{t^2 h}{1-vt-h^2t^2} \quad .$$

Note that $f(t)$, $g(t)$, and $F(t,h,v)$, $G(t,h,v)$ are **rational functions**. In particular, taking the **partial-fraction** decomposition of $f(t)$, over the reals one easily gets the **asymptotics** for $a(n)$, namely $(\sqrt{5}/\phi) \cdot \phi^n$, where ϕ is the Golden Ratio. Also by differentiating w.r.t. h and v and then plugging-in $h = 1$ and $v = 1$ one gets the generating function for "the total number of horizontal tiles" and "the total number of vertical tiles" respectively, from which once again, we can deduce asymptotics and get the asymptotic averages by dividing by the already-known asymptotic "number of tilings". Ditto for the higher moments and correlation (of course in this toy problem the correlation is tautologically -1).

But there is more! Now that we have quick ways for computing $a(n)$ and $b(n)$, up to very large n, either via the generating function or directly by $(MarkovianScheme)$, we can use the latter, with the help of **Wilf's methodology**[W] to do **sequencing**, **ranking**, and **random selection**. Let's just consider the last, that of generating **uniformly at random**, such a tiling of the $2 \times n$ rectangular board. Use a loaded coin, with probabilities $a(n-1)/a(n)$ and $b(n)/a(n)$ to decide whether to cover the leftmost-bottommost cell with a vertical or horizontal tile respectively. In the former case, continue recursively. In the latter case, use a very biased coin with probabilities

$b(n)/a(n-2) = 1$ and 0 to decide whether to tile the leftmost 0 at the top row, with a horizontal tile or vertical tile respectively, and then continue recursively.

The General Case

The same reasoning applies to an **arbitrary** (but fixed, i.e. numeric) width k and an *arbitrary* (but of course finite) set of (finite) "tiles". Here a tile is any finite set of lattice points, that does not need to be connected, and a tiling is a covering by translations of the tiles.

The logic is the same as above, but instead of only one 'auxiliary' set $\mathcal{B}(n)$ (that was trivially isomorphic to $\mathcal{A}(n-2)$) we get many more such 'stepping-stones'. These auxiliary sets are tilings of jagged rectangular boards that can be represented as, for example, (here the width, k, is 4) :

$$\begin{matrix} 0101\ldots0 \\ 1001\ldots0 \\ 0000\ldots0 \\ 1110\ldots0 \end{matrix} \quad ,$$

where ... stands for 0's (i.e. still unoccupied cells). We can code such jagged boards on the computer just as above as matrices of 0's and 1's leaving the ... implicit, and making it right-justified (in addition to left-justified), and hence there should be at least one 1 in the rightmost column. We also assume that the leftmost column has at least one 0, or else it is equivalent to a smaller board with that leftmost column removed (giving the removed column its due weight by multiplying by t when we do the equations). For each such jagged board we (or rather the computer) locates the **fundamental free cell** which is the lowest 0 in the leftmost column. We then consider which of the tiles in the set of tiles can cover that cell, and in what relative position of the considered tile. For each such scenario, the computer places that tile, getting a different jagged board, that may be a previously-encountered one or a brand-new one. For each new board, we do it again. The computer checks at each stage whether each and every jagged-board encountered so-far already has a structure theorem expressing it a a union of isomorphic copies of other jagged boards, and if not, keeps creating new structure theorems, that in turn usually bring in new kinds of jagged boards. Eventually there won't be any new ones (pigeonhole!) and then the process halts.

Next, to get generating functions, the computer (all by itself) takes weight-enumerators, getting a (usually very large) system of linear equations for the set of unknown weight-enumerators for these jagged boards, that includes, in its midst, the original, non-jagged rectangular board. Then, Maple `solves` this system, giving in particular the generating function for the original set, the rational function, (in t, or in t as well as in the tile-variables) whose Maclaurin series's coefficient of t^n is the number of (or weight-enumerating polynomial for) tilings of a $k \times n$ rectangular board by the given set of tiles.

Once the generating function is known, we can process it, using Maple's built-in partial-fraction (over the reals) (`parfrac`) and differentiation, `diff`, to get asymptotic size, and asymptotic averages, variance, covariance, and higher moments for the random variables "number of occurrences of a given tile-type".

Disclaimer: This last part involves floating-point calculations and may not be entirely rigorous. In principle, of course, one can stay within algebraic numbers, but the output would not be insightful to human eyes.

The Maple Package TILINGS

All this (and more!) is implemented in the Maple package `TILINGS`, that can be downloaded from

`http://www.math.rutgers.edu/~zeilberg/tokhniot/TILINGS` .

Once you downloaded it into a directory in your computer, calling it `TILINGS` (and **not** `TILINGS.txt`), go into Maple, and type:

`read TILINGS:`

and then follow the instructions given there. In particular, for a list of the main functions, type `ezra();` and for help with a specific function, type `ezra(FunctionName);`.

Also available is a precursor Maple package `RecTILINGS` that only handles rectangular tiles. It can be downloaded from:

`http://www.math.rutgers.edu/~zeilberg/tokhniot/RecTILINGS` .

A Very Quick Overview

Full details are given on-line, but for the benefit of the lazy reader who does not use Maple, let me just mention that the main functions are:

`GFt(k,Tiles,t)` and `GFtH(k,Tiles,t,H)` .

The former computes the generating function (in t only) and the latter the weight-enumerator (in t and in the variables $H[tile]$, one for each tile).

$Rt(a,b)$ is the package's shorthand for a $b \times a$ rectangular tile. For example, the toy problem we did above by hand is reproduced by:

`GFt(2,[Rt(1,2),Rt(2,1)],t);` and

`GFtH(2,[Rt(1,2),Rt(2,1)],t,H);` .

`Sidra(k,ListOfTiles,N);` gives you the first $N+1$ terms of the enumerating sequence (that can automatically be sent to Sloane's database), while `SidraW(k,H,ListOfTiles,N);` gives you the first $N+1$ terms in the sequence of weight-enumerators, that are polynomials in $H[tile]$'s.

`Astat(k,ListOfTiles,n);` gives you the asymptotic statistics: average number of occurrences for each tile, the variance, and the asymptotic covariance matrix for these random variables (in the order of appearance in the list `ListOfTiles`). `AstatV` is a verbose version of `Astat`.

Finally, `RandTiling(ListOfTiles,k0,n0);` gives you a (uniformly) random tiling of a $k0 \times n0$ rectangular board, using the given tiles (e.g. try: `RandTiling([Rt(1,2),Rt(2,1)],4,10);`), and

`RandTilingNice(ListOfTiles,k0,n0);` gives you the same in nice, matrix form.

`BuildTree` and `BuildTreeH` actually gives you the Markovian Enumeration Scheme, complete with generating functions, and `Sipur` and `SipurArokh` tells you everything you always wanted to know about the tilings.

Sample Input and Output

the webpage of this article,

`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/tilings.html` ,

contains several sample input and output files for `SipurArokh`. Readers can generate their own output for their favorite set of tiles and for their favorite width.

An "Almost" Automatic Proof of Kasteleyn's Formula

In the very special case of two dimer tiles (1×2 and 2×1), Kasteleyn and Fisher&Temperley ([K][FT], see [AS] for a beautiful recent survey) gave their deservedly celebrated beautiful formula for the weight-enumerator of an $m \times n$ rectangle for *arbitrary* (i.e. symbolic) m and n. To wit, if m and n are both even positive integers, and z and z' the variables for the horizontal and vertical dimers respectively (what we called H and V above), then

$$Z_{m,n}(z,z') = 2^{mn/2} \prod_{r=1}^{m/2} \prod_{s=1}^{n/2} \left[z^2 \cos^2 \frac{r\pi}{m+1} + z'^2 \cos^2 \frac{s\pi}{n+1} \right] \quad , \qquad (K)$$

with a similar formula when n is odd. At this time of writing, computers can't prove this formula in general. But for any *specific* m (but general(!) n), it is now a routine verification, thanks to our Maple package `TILINGS`. Of course, in practice, it can only be done for small values of m ($m \leq 10$ on our computer), but with bigger and future computers, one would be able to go further. The proof is a beautiful example of (rigorous!) generalization from finitely many cases, combined with 'general-nonsense' linear algebra *handwaving*, that is nevertheless fully rigorous.

Let's call the right side of (K) $W_{m,n}(z,z')$. We have to prove, for our given m, that $W_{m,n}(z,z') = Z_{m,n}(z,z')$ for *all* n. For a fixed even m it is readily seen that $W_{m,n}(z,z')$ is expressible as a product of $m/2$ different dilations of $U_n(z)$, the Tchebycheff polynomials of the second kind. It is well-known (and easily seen) that $U_n(z)$ satisfies a (homogeneous) *linear recurrence equation with constant (i.e. not depending on n) coefficients* of order 2, and hence do does any of its dilations $U_m(\beta z)$, and the product of $m/2$ of these creatures consequently satisfies a linear recurrence equation with constant coefficients of order $2^{m/2}$. Hence its generating function, in t, is rational function whose denominator has degree $2^{m/2}$ and numerator degree $< 2^{m/2}$. We have to prove that this generating function coincides with the 'real thing', the rational function outputted by procedure `GFtH` of `TILINGS`, that also turns out to have the same degrees. But to prove that two rational functions whose numerator degree is p and denominator degree is q are identical, we only have to check, by elementary linear algebra, that the first $p+q+2$ terms in their Maclaurin series coincide, and this is a routine check, done in our package `TILINGS` by typing `ProveKasteleyn(m);`, for the general case (with z and z'), and `ProveKasteleyn1(m);`, for the straight-enumeration case ($z = z' = 1$).

By hindsight, Kasteleyn and Fisher&Temperley were extremely lucky, since in the very special case of the two dimer tiles, they were able to use graph theory and Pfaffians. The general case seems, at present, out of reach. Even the monomer-dimer problem is still wide open, let alone an "explicit" expression for the weight-enumerators for tilings of an $m \times n$ rectangle using other sets of tiles. Our Maple package `TILINGS` is a *research tool* that can automatically discover and rigorously prove results for **specific** m but general n. Let's hope that the human can use it to discover new **ansatzes** by which to conjecture and hopefully prove some "explicit" form for the monomer-dimer and more general sets of tiles, where *both* m and n are **symbolic**. From this, one should be able to extract, at least, an "explicit" expression for the so-called *thermodynamic limit*, or at the very least, determine rigorously the *critical exponent*.

I hope to explore these speculations in a subsequent article.

References

[AS] Frederico Ardila and Richard S. Stanley, *Tilings*, preprint, `www.arxiv.org/math.CO/0501170`, 2005.

[FT] M. Fisher and H. Temperley, *Dimer Problems in Statistical Mechanics-an exact result*, Philos. Mag. **6** (1961), 1061-1063.

[K] P. W. Kasteleyn, *The statistics of dimers on a lattice: I. The number of dimer arrangements in a quadratic lattice*, Physica **27** (1961), 1209-1225.

[P] G. Polya, *On Picture Writing*, Amer. Math. Monthly **63** (1956), 689-697.

[W] Herbert S. Wilf, *A Unified Setting for Sequencing, Ranking, and Selection Algorithms for Combinatorial Objects*, Adv. in Math. **24** (1977), 281-291.