SOME IDEAS FOR THE FAT MAN PROBLEM

DAVID WHITE

We seek to solve a 2 dimensional analog of the Fat Man problem.

1. Motivating question in 1 dimension

Suppose there is a (linear) bar with n stools. This bar is frequented only by fat men. When a fat man sits down, no one can sit to his left or to his right. The fat men walk in one by one and each sits at a stool chosen uniformly at random from all stools, subject to the rule that he can't sit next to another fat man. Continue until the bar is full. What's the expected number of fat men who will be able to sit? This is related to the famous parking problem first studied by Renyi ([1]). A discussion can be found in [2].

Note: One might consider whether we first restrict the set of seats available to the fat man and choose then uniformly at random from those seats, or whether we let fat men choose uniformly at random from all seats and then make them choose again if they pick a seat which breaks the rule above. I am positive that the distribution of seats occupied at each time step is the same under either scheme, basically by the definition of the uniform distribution. Perhaps someone should write down an argument for this or perhaps it would be obvious to any reader.

Answer to the motivating question: Asymptotically, the expected number of patrons the bar fits is $\left(\frac{1}{2} - \frac{1}{2e^2}\right) * n$. This was proven in [2], along with a generalized version where the men are so fat that no one can sit within k seats of them. The proof uses the fact that each time a fat man sits down he splits the bar into two smaller bars, thereby giving a recursion to solve. Note that if the bar is circular then as soon as the first man sits down we're in the linear bar situation with 3 less seats. So that problem is also solved and is in fact no easier or harder than the linear bar.

The problem was also considered by Philippe Flajolet in [3], where he includes Maple code for simulations. I don't believe anyone has considered the problem where some men are fat ($k = 1$) , some are fatter ($k = 2$), some are even fatter ($k = 3$), etc. I imagine the framework outlined below could be used to prove limits in those cases without the need for the recursions used in [2].

2. 2 dimensional analog

This process can be done on any graph, and can be thought of as a way to choose a maximal independent set. Simply pick a vertex uniformly at random and put it in your pocket. Then remove all its neighbors. Repeat until there are no more vertices. You now have a maximal independent set in your pocket. In [4] the authors considered the problem on a $n \times 2$ ladder. This can be thought of as a dinner party with unfriendly guests, who don't want to sit next to each other or across from each other. As with the 1 dimensional case above, each time a guest sits down he disconnects the graph, leading to a recursion which can be solved. The algebra here is much messier, but the asymptotic limit (the expected proportion of the $2n$ seats you'll be able to use) turns out to be $\frac{1}{2} - \frac{1}{4e}$.

Date: December 17, 2011.

Also worth noting is that in [4] the authors show the limit exists for any $n \times m$ grid and they run simulations to make conjectures for small m. Already for $m = 3$ the method used so far fails because removing a vertex and it's neighbors need not disconnect the graph. Making the graph into a torus (identifying opposite walls) reduces the $m = 3$ case to the original ($m = 1$) problem. The $m = 5$ torus reduces to the $m = 3$ ladder, so moving to a torus doesn't really help, although it should not change the asymptotic answer. We'd like to come up with a method to prove the $m = 3$ case and hopefully it would generalize to also prove the $m = 4, 5, \ldots, n$ cases.

First attempt: noting that the graph does split when we remove a vertex on the middle rung of the $m = 3$ ladder, we thought about using some sort of exchangeability of events in a random process to try to force all such removals to occur first. We were advised by Michel Dekking that this would probably not work, as it would radically change the asymptotic limit. Another idea would be to tweak the process so at each time step you remove a set of points rather than an individual point. If the sets are chosen to be disconnecting this would give a recursion. We were told later by Mike Keane that Michael Aizenman tried this on a percolation problem and it didn't work.

Second attempt: Following Wormald's approach in [5] we thought we could look at a continuous version of this problem and represent it as a differential equation. By Wormald's theorem, solving this differential equation (might?) give the asymptotic limit. The problem here seems to be that physicists have already considered the continuous version of this problem under the name "the hard particle problem" and concluded that the differential equation was too hard to solve. They moved to the discrete problem on an $n \times m$ grid and it appears no one has solved it yet. It is possible some physicist has solved the $m = 2$ or $m = 3$ ladder; we haven't done an extensive search of the physics papers on this topic, but the interested reader could start such a search with the papers listed in the references ([8]-[15]). It might be worth re-doing the work in [4] using Wormald's approach.

Third attempt: Consider the problem as a subshift of finite type. We have hope this method will work. Ethan Coven recommends emailing Fabien Durand at the Université de Picardie with this idea.

3. Subshift of Finite Type in 1-dimension

In the 1 dimensional case we can think of the process of fat men entering the bar as a process to choose a length n string of 0s and 1s. Here a 0 will mean an empty seat and a 1 will mean an occupied seat. The fatness of the men is quantified by a rule saying we can never have 11 in the sequence. Insisting that the bar is full is quantified by a rule saying we can never have 000 in the sequence. It turns out that we now have a 2-3 code, much like IBM studied in the early years of personal computers. It's similar to the Golden Mean sequence, which forbids 11 but not 000.

To find all possible strings of length n we can use the theory of subshifts of finite type to write down an infinite sequence and then cut it down to one of length n. To do this, consider strings of length 2 and figure out all legal ways to add one more element: $00 \to 001, 01 \to 010, 10 \to 101$ or 100 (picked with uniform probability). If we then take these strings of length 3 and cut off the first number (this is the shift) we have a process which takes strings of length 2 to strings of length 2. The key information about this process is its entropy, which computes how fast the number of allowable configurations grows as $n \to \infty$. This information is contained in the following matrix M, where rows and columns are labeled by $00, 01, 10, 11$ and we put a 1 in the matrix if we can get from the row label to the column label by one step of the process:

$$\begin{array}{c|c} & \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \hline \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Because 11 is an illegal configuration we can remove those rows and columns from the above. M^n counts hte number of sequences of length n. The largest eigenvalue λ of M makes it easy to compute the number C_n of possible saturated configurations. For fixed n, $C_n = c * \lambda^n$ for some constant c. This is different from computing the number of 1s, but a twist (possible already known in the theory of subshifts of finite type) should allow you to just compute the number of 1s. Alternately, you could list all possible strings of length n and then to find the asymptotic limit we would simply have to compute the probability of reaching each configuration (e.g. by counting the number of ways to reach it), multiply this by the number of 1s, and take the sum of all those values. It is important to note that even though the fat man process and the sequence process produce the same list of saturated configurations, they might do so with different probabilities. It is unclear how this will affect the asymptotic limit, but the number of configurations above is an algebraic number (because λ is), so the asymptotic limit given by the sequence method should also be algebraic, unlike $\frac{1}{2} - \frac{1}{2e^2}$.

Here is a related question: if you pick a configuration uniformly at random from all possible saturated configurations (i.e. independent sets) then you should get a different expected number of 1s than after doing the fat man problem. It might be worthwhile to run an experiment to confirm this and see how far apart the two limits are. A common problem in the theory of subshifts of finite type is to find the configuration of maximum entropy, i.e. where all cylinder sets have the same probabilities. This would let you prove correct the experimental value above, but this value would certainly be different from the asymptotic limit for the fat man process. It would be interesting to know how different these values are, and perhaps to bound one by the other. Note that a cylinder set is by definition:

$$C_t[a_0, \cdots, a_m] = \{x \in \{0,1\}^{\mathbb{Z}} : x_t = a_0, \ldots, x_{t+m} = a_m\}$$

The theory of subshifts of finite type should allow you to compute the expected number of 1s if you were to pick saturated configurations with equal probability, but to get it to output the asymptotic limit achieved by the fat man problem you might need to change the transition probabilities, e.g. in the step $10 \to 00$ or 01 for the 1-dimensional version. It would be great to know if there was some way to set those probabilities so that the sequences produced by the subshift of finite type occurred with the same probabilities as those produced by the fat man process. This is a question about change of measure, since the measure on subshifts of finite type is based on cylinder sets. What we're asking is for the measure of each cylinder to equal the probability of reaching that configuration if you use the fat man process to distribute the 1s. I suspect this will be difficult, since each fat man entering the bar sees a global picture, so it'll be hard to govern his action by purely local rules without knowledge of the whole sequence. Each arrival to the bar changes the probabilities for each seat the next patron might take, and it's unclear how to capture this in a transition probability. Perhaps the solution is to choose Markovian transition probabilities.

It might also be interesting to study how to get from one configuration to another using local rules. This originally arose as a way to make the process irreducible but it seems related to partition functions in statistical physics and so might be interesting on its own, without the context of the fat man problem. Michel Dekking thinks that these local rules would also give rise to a subshift of finite type.

4. Subshifts of finite type for ladder

To use the method of subshifts of finite type on the $n \times 3$ fat man ladder we need a matrix whose rows and columns are labeled by blocks of size 2×3. It should not be hard to write down all allowable configurations, though they may also need to depend on possibilities for where fat men are sitting in columns on either side of the two in question. Because this ladder is still close to a 1-dimension problem, some of the theory above still holds, i.e. you can compute the largest eigenvalue of and use it to count configurations. If the theory works out nicely in the 1-dimensional case above then we might get a proof of the conjectured value for the $n \times 3$ ladder (or at least a bound on it). Because this proof would not depend on a recursion it could probably be generalized to an $n \times m$ ladder for fixed m.

Finding the asymptotic limit on $\mathbb{Z}^2$ with this method seems much harder, since the theory of 2-dimensional subshifts of finite type is less well-developed. Finding the expected number of configurations without two 1s next to each other is the hard particle problem, which Mike Keane has been wondering about for 40 years. Mike suggests working on the honeycomb lattice (i.e. chickenwire) but thinks this will still be extremely difficult. The answer will have corollaries in tiling theory and a number of other hard areas. Note that Freeman Dyson worked out the Golden Mean sequence in $\mathbb{Z}^2$ and it was much harder than in $\mathbb{Z}$. One could try to follow his work to solve the $m, n \to \infty$ case of the Fat Man problem, but I have no idea if there is hope in this approach. It seems that doing this for fixed m first is a much better idea.

5. References

(1) A. Rényi, "On a one-dimensional problem concerning random space-filling", *Publications of the Mathematical Institute of the Hungarian Academy of Sciences 3* (1958) 109127.

(2) H. D. Friedman, D. Rothman, "Solution to: An unfriendly seating arrangement (problem 62-3)", *SIAM Review 6 (2)* (1964) 180182.

(3) P. Flajolet, "A Seating Arrangement Problem," http://algo.inria.fr/libraries/autocomb/fatmen-html/fatmen.html

(4) K. Georgiou, E. Kranakis, and D. Krizanc, "Random maximal independent sets and the unfriendly theater seating arrangement problem" *Discrete Mathematics* (2009) 5120-5129

(5) N.C. Wormald, "The differential equation method for random graph processes and greedy algorithms", in M. Karoński and H.J. Prömel, editors, *Lectures on Approximation and Randomized Algorithms*, pages 73155. PWN, Warsaw, 1999.

(6) B. Marcus and D. Lind, *An Introduction to Symbolic Dynamics and Coding* Cambridge University Press, 1995.

(7) N.C. Wormald, "Differential equations for random processes and random graphs", *Annals of Applied Probability 5* (1995), 1217-1235.

(8) R. J. Baxter, "Planar lattice gases with nearest-neighbour exclusion", *Annals of Combinatorics 3* (1999) 191203.

(9) J. Bouttier, P. D. Francesco, E. Guitte, "Critical and tricritical hard objects on bicolorable random lattices: Exact solutions", *J.Phys. A35* (2002) 38213854 Also available as arXiv:cond-mat/0201213.

(10) J. Bouttier, P. D. Francesco, E. Guitte, "Combinatorics of hard particles on planar graphs", *J.Phys. A38* (2005) 45294559 Also available as arXiv:math/0501344v2.

(11) S. R. Finch, "Several Constants Arising in Statistical Mechanics", *Annals of Combinatorics* (1999) 323335.

(12) F. H. Stillinger, H. Sakai, S. Torquato, "Critical and tricritical hard objects on bicolorable random lattices: Exact solutions", *Physical Review E67* (2003) 031107, 110.

(13) S. Torquato, A. Donev, F. Stillinger, "Breakdown of elasticity theory for jammed hard-particle packings: Conical nonlinear constitutive theory", *International Journal of Solids and Structures 40* (2003) 71437153.

(14) J. K. MacKenzie, "Sequential filling of a line by intervals placed at random and its application to linear adsorption", *The Journal of Chemical Physics 37 (4)* (1962) 723728.

(15) N. J. Calkin, H. S. Wilf, "The number of independent sets in a grid graph," *SIAM Journal of Discrete Mathematics 11 (1)* (1998) 5460.