## Hall's Marriage Lemma

Jonathan L.F. King
*University of Florida, Gainesville FL 32611-2082, USA*
squash@ufl.edu
Webpage http://squash.1gainesville.com/
2 April, 2018 (at *08:30*)

**Entrance.** We start with the classic thm of Philip Hal. We have sets $\mathfrak{B}$ and $\mathcal{G}$ [“boys” and “girls”, possibly infinite] and a bipartite graph $\Gamma = ((\mathfrak{B}, \mathcal{G}), \mathsf{E})$. We write $b\mathsf{E}g$ if boy $b$ and girl $g$ know each other. Let $\female(b)$ be the *set* of girls known by $b$. And for a set $B \subset \mathfrak{B}$, use

$$\female(B) \quad := \quad \bigcup_{b\in B} \female(b)$$

for the girls known by at least one $B$-boy. Analogously, use $\male(g)$ and $\male(G)$ for boys known by girls.

A “***marriage*** for the boys” is an *injection* $f{:}\mathfrak{B} \hookrightarrow \mathcal{G}$ st. for each $b$, we have $b\mathsf{E}f(b)$.

The ***Hall condition*** on $\Gamma$ is:

HC: *Each* $B \subset \mathfrak{B}$ *has* $|\female(B)| \geqslant |B|$.

Evidently HC is a necessary condition for a marriage.

**1: Marriage lemma** (Philip Hall, 1935). *Suppose bipartite graph* $\Gamma = ((\mathfrak{B}, \mathcal{G}), \mathsf{E})$ *has* $\mathfrak{B}$ *finite. Then there is a marriage for the boys* IFF $\Gamma$ *satisfies HC.* ♢

1a: *Remark.* When $\mathfrak{B}$ is infinite, HC does not imply a marriage. Consider boys $\mathbb{N}$ and girls $\mathbb{Z}_+$. *Boy* $b_0$ *knows all the girls, and each other* $b_n$ *knows only* $g_n$. Each *proper* subset $B \subset \mathbb{N}$ can be married-off: Pick $K \in \mathbb{N} \smallsetminus B$ and marry $b_0$ to $g_K$. For the remaining $n \in B$, marry $b_n$ to $g_n$.

OTOHand, we can't marry-off *all* the boys; the wife $g_K$ of $b_0$ leaves poor $b_K$ with no-one to marry.

The below proof uses induction on $|\mathfrak{B}|$, doing divorces to marry-in the new boy. The above CEX shows that there can*not* be an induction proof-by-extension; the divorces are necessary, even with look-ahead. □

*Pf of* (1). Suppose we have married-off finite set $\mathfrak{B}$ into [possibly infinite] $\mathcal{G}$. We have a new boy $b_0 \notin \mathfrak{B}$ whom we wish to marry-off. Our goal is to find a ***chain***

$$*: \quad b_0 \to g_1 \twoheadrightarrow b_1 \to g_2 \twoheadrightarrow b_2 \to \ldots \to g_{K-1} \twoheadrightarrow b_{K-1} \to g_K\,,$$

where: *Girl* $g_K$ *is unmarried, each other* $g_n$ *is married to* $b_n$*, and each* $b_{j-1}$ *knows* $g_j$. Divorce these married-girls, then marry each $b_{j-1}$ to $g_j$. Now all boys in $\mathfrak{B} \sqcup \{b_0\}$ are married.

**Producing a chain.** “Mark” $b_0$. Iteratively mark additional girls and boys as follows:

- Mark each girl known by a marked boy.
- Mark each boy married to a marked girl.

This process must eventually stabilize, as $\mathfrak{B}$ is finite. At this point, let $B$ and $G$ be the sets of marked boys and girls. By defn $G = \female(B)$, so the Hall condition says $|G| \geqslant |B|$.

Each $B$-boy *except* $b_0$ is married; so $G$ has precisely $|B|-1$ wives. Thus *there is some unmarried* $G$*-girl.* Pick one. In the marking-process, she was introduced at some stage, $K$. Hence she is $g_K$ of some $(*)$-chain. ♦

## Distinct-cards Problem

For the cards in a playing-deck, denote the ranks A,2,. . . ,J,Q,K by $r_1, r_2, \ldots, r_{13}$.

**2: Distinct-cards thm.** *Deal a randomized deck into 13 piles of four cards apiece. Now* <u>*remove*</u> *some three cards. Then it is always possible to choose one card-per-pile so that all 13 ranks were chosen.* ♢

*Proof.* Imagine that each pile of cards is on its own little tray. The trays are the “boys”, the ranks are the “girls” and the cards are the $52-3 = 49$ edges of the bipartite graph. Does this graph satisfy Philip Hall's condition?

In a set, $\mathcal{C}$, of $n$ many cards, no rank occurs on more than 4 cards, so the number of ranks occurring in $\mathcal{C}$ is at least $\lceil n/4 \rceil$. A collection of $K$ many trays has at least $n := 4K-3$ many cards, so this collection has at least $\left\lceil \frac{[4K]-3}{4} \right\rceil \stackrel{\text{note}}{=\!=} K$ many ranks. I.e, each set of $K$ boys “knows” at least $K$ many girls. ♦

## Land matching

3: The Hunter/Farmer problem. *There is an island which, from time immemorial, has been divided into $N$ equal-area farming regions, taking up the whole island. It is also divided into $N$ equal-area hunting tracts, taking up the whole island.*

*There are $N$ married couples on the island; the wives hunt and the husbands farm. We would like to be able to assign tracts to wives and farms to husbands so that each couple could build a house on territory common to both. Indeed, territory with at least area $\boxed{\delta_N \cdot \text{Area}(\textit{Island})}$. Determine the largest $\delta = \delta_N$ which that works for every division of the island into tracts/regions.* ♢

*H/F answer.* Normalize by Area(Island) $= N$. We'll describe our maximal $\delta_N$ in terms of $\varepsilon_N := N \cdot \delta_N$, using quadratic $G(x) = G_N(x) := [[N+1] - x] \cdot x$. We'll show that

3a:
$$\text{If } N \text{ odd: } \varepsilon_N = 1/G(\tfrac{N+1}{2}) \stackrel{\text{note}}{=\!=} \frac{4}{N^2 + 2N + 1};$$
$$\text{If } N \text{ even: } \varepsilon_N = 1/G(\tfrac{N}{2}) \stackrel{\text{note}}{=\!=} \frac{4}{N^2 + 2N}.$$

The Marriage lemma will show (3a) sufficient, and an example will show (3a) necessary. □

*Preliminaries.* For an $\varepsilon > 0$, say that a tract $T$ ***knows*** farm $F$ if
$$\text{Area}(T \cap F) \geqslant \varepsilon.$$

Let $\mathbf{H}$ be the union of some $\mathbf{h} \in [1 \,..\, N]$ many tracts. Let $\mathbf{k}(\mathbf{H}, \varepsilon)$ denote the # of **"*good*"** farms which know at least one $\mathbf{H}$-tract. Then *"Hall's for $\mathbf{H}$"* is

$C_{\mathbf{H}}$:
$$\mathbf{k}(\mathbf{H}, \varepsilon) \geqslant \mathbf{h}.$$

And Hall's condition is that $(C_{\mathbf{H}})$ holds for *every* $\mathbf{H}$. □

Sufficiency of (3a). Consider $\mathbf{h}$ and $\mathbf{H}$ as above. Then $\mathbf{k} = \mathbf{k}(\mathbf{H}, \varepsilon)$ is the number of $\mathbf{H}$-good farms; call the other farms **"*bad*"**. Condition $(C_{\mathbf{H}})$ is non-trivial only when $\mathbf{h} > 0$ and $\mathbf{k} < N$. Since $\mathbf{h}$ is positive,

†: *Each bad farm intersects $\mathbf{H}$ in area strictly less than $\mathbf{h}\varepsilon$.*

And since $N - \mathbf{k}$ is positive, our (†) gives us the *strict* inequality

‡:
$$\mathbf{h} = \text{Area}(\mathbf{H}) < \mathbf{k} \cdot \overbrace{1}^{\text{Good}} + [N - \mathbf{k}] \cdot \overbrace{\mathbf{h}\varepsilon}^{\text{Bad}}$$
$$\stackrel{\text{note}}{=\!=} N\mathbf{h}\varepsilon + \mathbf{k}[1 - \mathbf{h}\varepsilon].$$

**When is $\varepsilon$ small enough for $(C_{\mathbf{H}})$?** Suppose $\varepsilon$ is so small that $\mathbf{k} := \mathbf{h} - 1$ *fails* (‡), forcing $\mathbf{k}$ to be *at least* $\mathbf{h}$. *Failure* of (‡) is equivalent to

$$\mathbf{h} \geqslant N\mathbf{h}\varepsilon + [\mathbf{h} - 1][1 - \mathbf{h}\varepsilon]. \qquad \textit{Rewriting,}$$

$*$:
$$1 \geqslant \varepsilon \cdot [N\mathbf{h} + \mathbf{h} - \mathbf{h}^2] \stackrel{\text{note}}{=\!=} \varepsilon \cdot G_N(\mathbf{h}).$$

The largest $\varepsilon$ satisfying $(*)$ for *every* $\mathbf{h} \in [1 \,..\, N]$ is thus

3b:
$$\varepsilon_N := \operatorname*{Min}_{\mathbf{h} \in [1 \,..\, N]} \frac{1}{G(\mathbf{h})} = 1 \Big/ \operatorname*{Max}_{\mathbf{h} \in [1 \,..\, N]} G(\mathbf{h}).$$

Parabola $\mathbf{h} \mapsto G(\mathbf{h})$ has its unique maximum at real number $\mathbf{h} = \frac{N+1}{2}$. Since $G()$ is strictly convex-down and symmetric, the $G$-maximum over $\mathbb{Z}$ is obtained at the/an integer which is closest to $\frac{N+1}{2}$. Hence (3a).♦

Necessity of (3a). (27 Jan 1995) The island is a horizontal row of $N$ many $1 \times 1$ tracts. Let $\mathbf{H}$ be the union of the leftmost $\mathbf{h}$ tracts.

Divide $\mathbf{H}$, by horizontal lines, into $\mathbf{h}$ many rectangles. Define the bottom-most $\mathbf{k} := \mathbf{h} - 1$ of these rectangles to be farms.

**The Other farms.** Our "bottom-most"-farms leave an $\mathbf{h} \times \frac{1}{\mathbf{h}}$ ***Panhandle*** of $\mathbf{H}$ uncovered. Define the $[N - \mathbf{k}]$ "other" farms to equi-divide the Panhandle into horizontal rectangles, and then apportion the rest of ISLAND $\smallsetminus$ $\mathbf{H}$ to the "other"-farms in any way, so that every farm has area 1.

The tracts of $\mathbf{H}$ equi-split the Panhandle vertically, whilst the "other"-farms equi-split the Panhandle horizontally. Hence,

> *Each $\mathbf{H}$-tract $T$ has its part in the Panhandle equi-split by the set of "other"-farms.*

There are $[N - \mathbf{k}]$ "other"-farms, so the intersection of an "other"-farm $F$ with an $\mathbf{H}$-tract $T$ has

$$\text{Area}(T \cap F) = \frac{1}{N - \mathbf{k}} \cdot \text{Area}(T \cap \text{Panhandle})$$
$$= \frac{1}{N - [\mathbf{h} - 1]} \cdot \frac{1}{\mathbf{h}} \stackrel{\text{note}}{=\!=} 1 \big/ G_N(\mathbf{h}).$$

If $\varepsilon$ exceeded this $1/G_N(\mathbf{h})$, then none of the "other"-farms know any of the $\mathbf{H}$-tracts, so $\mathbf{k}(\mathbf{H}, \varepsilon) \leqslant \mathbf{h} - 1$, violating $(C_{\mathbf{H}})$. Hence (3b). ♦

**Harry's request.** With Area(Island)$=N$, Harry asked for an encoding of the above example into an $N\times N$ doubly-stochastic matrix, $\mathsf{M}_{N,\mathbf{h}}$, where the entry in row $T$ and column $F$ is $\text{Area}(T\cap F)$. [So the sum of *all* the M-entries is $N$, the area of the island.]

Fix $\mathbf{h}\in[1\,..\,N]$, set $\mathbf{k}:=\mathbf{h}-1$, let the top $\mathbf{h}$ rows correspond to the tracts forming $\mathbf{H}$; the leftmost $\mathbf{k}$ cols correspond to the farms that are subsets of $\mathbf{H}$. Then

$$\text{3c:}\qquad \mathsf{M}_{N,\mathbf{h}} = \begin{bmatrix} \mathsf{A}_{\mathbf{h}\times\mathbf{k}} & \mathsf{B}_{\mathbf{h}\times[N-\mathbf{k}]} \\ 0_{[N-\mathbf{h}]\times\mathbf{k}} & \mathsf{D}_{[N-\mathbf{h}]\times[N-\mathbf{k}]} \end{bmatrix},$$

where $\mathsf{A},\mathsf{B},\mathsf{D}$ are constant matrices, with respective entries $\frac{1}{\mathbf{h}}$, $\frac{1}{\mathbf{h}\cdot[N-\mathbf{k}]}$ and $\frac{1}{N-\mathbf{k}}$. [Matrix D need not be constant; D just need be non-negative, with every M-row and M-col passing through it, having sum=1.]

*Symmetry* $\mathbf{h}\leftrightarrow[N-\mathbf{k}]$. The upper-constraint on $\varepsilon$ is the B-value, $\frac{1}{\mathbf{h}\cdot[N-\mathbf{k}]} \overset{\text{note}}{=\!=} 1/G_N(\mathbf{h})$.

Note that $\mathbf{h}+[N-\mathbf{k}] = N+1$, and this eqn is symmetric in $\mathbf{h}$ and $[N-\mathbf{k}]$. Thus, if we always make the above D constant, then *exchanging* the values of $\mathbf{h}$ and $[N-\mathbf{k}]$, changes $\mathsf{M}_{N,\mathbf{h}}$ into its *transpose*. So in the above"Necessity of (3a)" paragraph, rather than consider *all* $\mathbf{h}\in[1\,..\,N]$, we only needed the contraints from those $\mathbf{h}\in[1\,..\,\frac{N+1}{2}]$. [Of course, since we need to maximize $G_N$, we only need consider one integer $\mathbf{h}$ closest to $\frac{N+1}{2}$.] □