Hall’s Marriage Theorem and Pawn Diagrams

Colin McDonagh

Abstract—This paper is concerned with unreachable pawn diagrams and the subset of which can be generated using Hall’s Marriage Theorem. The result is 1 in 23 diagrams are unreachable by applying the theorem.

Index Terms—chess, combinatorics

I. INTRODUCTION

A. Background

A chess diagram [1] is defined as the contents of the 64 squares, whereas a position also takes into account which side is to move, castling rights and any en-passant square.

Certain single-side pawn diagrams are unreachable as pawns may only move forward or diagonally forward. Consider “Fig. 1”. In order for a pawn to be on a3, it must have come from b2. However, since there’s a pawn on b2, this diagram could never arise in a game.

![Fig. 1. A simple unreachable diagram](image)

This paper assumes pawns can move diagonally forward without limit. In reality pawns can only move diagonally by capturing an opposition chessman (a pawn or a piece), apart from the King.

B. Literature Review

There doesn’t seem to be many papers written on the number of diagrams and positions in chess. However, one particular paper [2] gives a detailed method for calculating an upper bound for the number of reachable positions in chess. Said paper mentions that it over-counts the number of positions because of unreachable pawn diagrams such as in “Fig. 1”, which is the motivation for this paper.

II. FILE-RANGES

A pawn on any valid square (a2-h7) maps to a range of files from which it could have started on. For example in “Fig. 2”, the pawn on d4 could have started on the b-f files. Let this range of files be known as a pawn’s range or file-range.

If either of the edge files (a, h) are included in a pawn’s range, and the size of the range is greater than 1, i.e. the pawn is off its starting square, then the file-range does not uniquely identify the pawn. Consider “Fig. 3”, where a6, b5 and c4 have the same file-range of a-e.

![Fig. 2. File-range for a pawn on d4. Line illustrates range.](image)

![Fig. 3. File-range which maps multiple pawn-squares.](image)

It’s interesting to note that a file-range which does not include an edge file must have an even length, as every square a pawn advances, its range increases by one either side of the pawn. This can be seen in “Fig. 2”. Therefore a file-range of b-c is invalid. A file-range which includes an edge file can have an odd size, as shown in “Fig. 3”.

III. THE UNREACHABLE CONDITION

A. Hall’s Marriage Theorem

Digressing shortly, [3] let S be a finite family of finite sets. Suppose that for every subfamily R of sets in S, the number of subsets in R is less than or equal to the total number of elements in those subsets:

$$|R| \leq \bigcup_{X \in R} X$$


This is known as Hall’s Marriage Theorem (which in this paper is also referred to as Hall’s Condition). A less formal definition exists which is where the theorem gets its name from: suppose there are a finite set of single men, and, for each man, a finite set of compatible women whom the man is attracted to. For any subset of men, if the number of men is greater than the size of the union of the mens’ sets of compatible women, then each man cannot find or be paired with a compatible woman.

B. Applying Hall’s Condition

If we substitute man for pawn then we have the following: suppose there are a finite set of pawns, and, for each pawn, a finite set of files from which the pawn could have started on. For any subset of pawns, if the number of pawns is greater than the number of files from which the pawns could have started on, then the diagram is unreachable. And hence Hall’s Condition can be used to verify whether a diagram is unreachable.

Buckets are a useful construct because the number of ways in which a Bucket can satisfy Hall’s Condition corresponds directly to a number of unreachable diagrams.

V. Multiple Buckets

A. Non-disjoint Buckets

An unreachable diagram may correspond to multiple Buckets for which Hall’s Condition is satisfied. If the Buckets are disjoint, this is trivial problem to solve. However, when non-disjoint, determining the number of ways in which Hall’s Condition can be satisfied for each Bucket is non-trivial.

Given a collection of Buckets, $S$, and for some subset $s \subseteq S$, let the intersection $I_s$ represent the squares common to all Buckets in $s$. Consider “Fig. 6” as an example.

B. Mapping Squares

Let the set of $k$-element subsets of $S$ be represented by $\binom{S}{k}$. Furthermore, let us define the prime of $s$ as:

$$s' = \left( \binom{S}{|s|} \right), s' \neq s$$

Using the intersection $s$ and the other $|s|$-fold intersections we can then define the intersection difference of $s$ to be:

$$D_s = I_s \setminus \bigcup_{q \in s'} I_q$$

In this way, each square covered by the Buckets is distinctly mapped to some $D_s$, as shown in “Fig. 7”.

1For $|s| = 1$, $I_s$ represents the squares covered by the single Bucket. This helps simplify “Algorithm 1”.

IV. File-Range Buckets

A. Definition of a Bucket

Let the set of file-ranges which are subranges of some range $r$, be defined as the File-Range Bucket (or just Bucket) of $r$, where a file-range $f$ is a subrange of $r$ if $f \subseteq r$. More directly put:

$$B_r = \{ f : f \in r \}$$

where $B_r$ is the Bucket of $r$.

For example, consider “Fig. 4” and “Fig. 5”.

B. Mapping Squares

Let the set of $k$-element subsets of $S$ be represented by $\binom{S}{k}$. Furthermore, let us define the prime of $s$ as:

$$s' = \left( \binom{S}{|s|} \right), s' \neq s$$

Using the intersection $s$ and the other $|s|$-fold intersections we can then define the intersection difference of $s$ to be:

$$D_s = I_s \setminus \bigcup_{q \in s'} I_q$$

In this way, each square covered by the Buckets is distinctly mapped to some $D_s$, as shown in “Fig. 7”.

1For $|s| = 1$, $I_s$ represents the squares covered by the single Bucket. This helps simplify “Algorithm 1”.

Fig. 5. Squares covered by the $a-c$ Bucket.

Fig. 6. Buckets $a-b$ (1), $a-c$ (2) and $b-f$ (3). $I_{1,2}$ is marked with circles, $I_{1,3}$ is marked with a square and $I_{2,3}$ is marked with small disks.

Fig. 7. Squares covered by the $b-f$ Bucket.
Fig. 7. Buckets a-b (1), a-c (2) and b-f (3). D_{1,2} is marked with a circle, |D_{1,3}| = 0, and D_{2,3} is marked with a small disk. Furthermore, D_{1,3} is marked with a gray X, D_{2} is marked with gray squares and D_{3} is marked with gray plus signs.

C. Counting Diagrams

Hall’s Condition must hold for each Bucket. Given Buckets \( B_1, B_2, \ldots, B_n \), and respective Bucket widths \( W_1, W_2, \ldots, W_n \), the set of unreachable diagrams for the Buckets can be represented by:

\[
U = \{ u : u \subseteq B_1 \cup B_2 \cup \ldots B_n \land |u \cap B_i| > W_i \}
\]

In order to count \( U \), we can deal with the given Buckets’ intersection differences, \( D \), which are easier to reason about on account of being disjoint.

Let \( P \) represent a counter for a number of pawns contained by a difference intersection \( d \), and let \( P \) be associated with \( d \) through dot notation, i.e. \( d.P \). The algorithm in “Algorithm 1” then describes how to traverse \( D \) to count the Buckets’ unreachable diagrams.

VI. Extending Buckets

Many unreachable diagrams consist of an unreachable grouping of pawns and some other reachable subset. For example, in “Fig. 8”, the a-c Bucket may be combined with a selection of 0-5 pawns on the light-gray highlighted squares to generate other unreachable diagrams.

One such unreachable diagram for the a-c Bucket is given in “Fig. 9”.

Algorithm 1 Counting Buckets’ unreachable diagrams

\[
\text{index} \leftarrow 0 \\
\text{while} \ \text{index} \geq 0 \ \text{do} \\
\quad \text{err} \leftarrow \text{SETMINIMUMS(index)} \\
\quad \text{if} \ \text{!err} \ \text{then} \\
\quad\quad \text{COMPUTEPERMUTATIONS} \\
\quad\quad \text{index} \leftarrow \text{INCREMENTD} \\
\quad \text{function SETMINIMUMS(index)} \\
\quad\quad \text{numPawns} \leftarrow 0 \\
\quad\quad \text{for} \ d \in D \ \text{do} \\
\quad\quad\quad \text{d.P} \leftarrow \text{GETMINIMUM}(d) \\
\quad\quad\quad \text{numPawns} \leftarrow \text{numPawns} + \text{d.P} \\
\quad\quad \text{return} \ \text{numPawns} > 8 \\
\quad \text{function GETMINIMUM(d)} \\
\quad\quad \text{dmin} \leftarrow \text{ldl} \\
\quad\quad \text{for} \ b \in \text{Buckets which contain} \ d \ \text{do} \\
\quad\quad\quad \text{bmin} \leftarrow \text{minimum number of pawns} \ b \ \text{needs from} \ d \ \text{such that it may fulfil Hall’s Condition by the end of the setMinimums traversal} \\
\quad\quad\quad \text{dmin} \leftarrow \text{min(dmin, bmin)} \\
\quad\quad \text{return} \ \text{dmin} \\
\quad \text{function COMPUTEPERMUTATIONS} \\
\quad\quad \text{permutations} \leftarrow 0 \\
\quad\quad \text{for} \ d \in D \ \text{do} \\
\quad\quad\quad \text{permutations} \leftarrow \text{permutations} \ast \left( \binom{|d|}{d.P} \right) \\
\quad\quad \text{store permutations} \\
\quad \text{function INCREMENTD} \\
\quad\quad \text{for} \ i = |D| - 1 \ \text{to} \ 0 \ \text{do} \\
\quad\quad\quad \text{if} \ |D_i| > D_i.P \ \text{then} \\
\quad\quad\quad\quad \text{D}_i.P \leftarrow D_i.P + 1 \\
\quad\quad\quad \text{return} \ i \\
\quad \text{return} \ -1
\]

Fig. 9. One unreachable diagram for the a-c Bucket.

VII. Duplicate Diagrams

A. An Example

Any two Buckets (Buckets and extended Buckets, but for brevity sake we’ll just refer to them as Buckets) may have a subset of common unreachable diagrams. For example, firstly consider “Fig. 10”, in which we see an unreachable
diagram for the a-c Bucket. Now consider the same diagram in “Fig. 11”, which illustrates that the diagram also belongs to the a-d Bucket.

**Fig. 10.** Another unreachable diagram for the a-c Bucket.

**Fig. 11.** The same unreachable diagram also belongs to the a-d Bucket.

**B. Applying the Principle of Inclusion and Exclusion**

We can account for duplicate diagrams using the Principle of Inclusion and Exclusion (PIE) sieve method. Let \( C \) be a collection of Buckets. Let \( I_k \) represent the set of \( k \)-fold intersections of members of \( C \) (duplicate diagrams which belong to \( k \) different Buckets in \( C \)). The general form of PIE \([5]\) states that:

\[
\left| \bigcup_{i=1}^{N} A_i \right| = \sum_{j=1}^{N} \left( (-1)^{j+1} \sum_{S \in \ell_j} |S| \right)
\]

In word form, PIE can count the number of distinct unreachable diagrams by first summing the number of unreachable diagrams which belong to any single Bucket, then subtracting the number of unreachable diagrams which belong to any 2 Buckets, then adding the number of unreachable diagrams which belong to any 3 Buckets etc.

**VIII. Results**

The results obtained are shown in “Table. 1".

From this table, the average percentage of all conceivable single-side pawn diagrams which are unreachable is 4.322%