# The Mutilated Chessboard Theorem in Z

R.D. Arthan
Lemma 1 Ltd.
rda@lemma-one.com

10 December 2006

**Abstract**

This note presents a statement and proof of the mutilated chessboard theorem in Z. The formulation is along the lines of the one proposed by McCarthy for a 'heavy duty set theory" theorem prover. The theorem and its proof are presented as a Z specification and a series of Z conjectures all of which have been mechanically verified using the ProofPower system. For the entertainment of the theorem-proving community, we conclude by proposing a rather harder challenge problem.

# 1 INTRODUCTION

The mutilated chessboard theorem was proposed over 40 years ago by John McCarthy as a "tough nut to crack" for automated reasoning [6]. The theorem is very simple: if a chess board is "mutilated" by removing two diagonally opposite corner squares, then the result cannot be tiled with domino-shaped tiles each covering two adjacent squares of the board. Why is this so? Well, as each domino covers one white square and one black, any such tiling must cover exactly as many black squares as it does white. However, the two diagonally opposite corner squares have the same colour, and so the mutilated board contains two more squares of one colour than it does of the other.

A discussion of why this theorem is a challenge for theorem-proving is outside the scope of this note. However, as McCarthy [5] points out, the counting argument given above should be "admitted in any heavy duty set theory", i.e., any general purpose theorem-proving system allowing reasoning about sets should be able to handle the proof. The theorem has indeed been proved in a number of systems including NQTHM [10], Mizar [2, 8], Isabelle [7].

The Z language [4, 9] provides a convenient notation for set theory. The ProofPower system [1] supports specification and proof in Z. This note presents a formalisation of the mutilated chess board theorem theorem in Z, prepared and checked using ProofPower. The treatment is in the spirit of [5], which turns out to admit a very concise rendition in Z.

$$\boldsymbol{mutilated\_chessboard\_thm} \;?\vdash \neg(\exists z{:}\ Tiling \bullet \bigcup z = MutilatedBoard)$$

Figure 1: Statement of the Theorem

$$\boldsymbol{Adjacent} \mathrel{\widehat{=}} \{x1,\ x2 : \mathbb{Z} \times \mathbb{Z} \mid abs(x1.1 - x2.1) + abs(x1.2 - x2.2) = 1\}$$

$$\boldsymbol{Domino} \mathrel{\widehat{=}} \{x{:}\mathbb{F}(\mathbb{Z} \times \mathbb{Z}) \mid \#x = 2 \wedge (\forall x1,\ x2{:}x \mid x1 \neq x2 \bullet (x1,\ x2) \in Adjacent)\}$$

$$\boldsymbol{Tiling} \mathrel{\widehat{=}} \{z{:}\mathbb{P}(\mathbb{P}(\mathbb{Z} \times \mathbb{Z})) \mid z \subseteq Domino \wedge (\forall x,\ y{:}z \bullet x = y \vee x \cap y = \{\})\}$$

$$\boldsymbol{Board} \mathrel{\widehat{=}} (0 \mathrel{..} 7) \times (0 \mathrel{..} 7)$$

$$\boldsymbol{MutilatedBoard} \mathrel{\widehat{=}} Board \setminus \{(0{,}0),\ (7{,}7)\}$$

Figure 2: The Specification

# 2 STATEMENT OF THE THEOREM

Our statement of the theorem is the Z conjecture shown in figure 1. It involves the sets, *Tiling* and *MutilatedBoard* defined along with the sets *Adjacent*, *Domino* and *Board* in the little Z specification shown in figure 2. The specification is based on that of [5] with some simplifications. Let us discuss each of the 5 sets in the order of their introduction in figure 2.

We represent a square on the chessboard as a pair of integers. Two squares are adjacent iff. the distance between them is 1 (where we can conveniently measure distances using the Manhattan taxi-cab metric). The set *Adjacent* represents this relation.

A domino in some position on the board is represented by the set comprising the two adjacent squares it occupies. The set of all such sets is our set *Domino*. Unlike McCarthy, we do not constrain our dominoes to lie on the chessboard, this makes our statement a little more general than his. (Yes, it is obvious that a tiling that contains dominoes that are not on the board doesn't tile the board, but the whole thing is "obvious" so we may as well formalise it as generally as we conveniently can.)

A tiling is a set of dominoes positioned so that no two dominoes overlap. Our set *Tiling* is the set of all such tilings, specified using the obvious and direct translation of that description into set theory. Paulson [7] argues that it is good to specify this set using induction. Certainly an inductive definition will be intuitive to someone whose mode of thinking about such things is to concentrate on the processes whereby they are constructed. The resulting definition in Z would describe the set *Tiling* as the smallest set of domino configurations containing the empty configuration and closed under the operation that adds a non-overlapping domino to a configuration. This is less general (since it does not allow for infinite tilings) and gives a bias towards an inductive proof, even though there is no induction in the informal proof.

As reflected in the definition of the set *Board*, the complete board is to be represented by the set of squares with co-ordinates in the range 0 to 7, which we can write as the cartesian product of the integer interval $0 \mathrel{..} 7$ with itself. The set *MutilatedBoard* is then obtained by removing the squares $(0, 0)$ and $(7, 7)$.

With our representation of dominoes, the set of squares covered by a tiling $z$ is the union $\bigcup z$. So, returning to the conjecture shown in figure 1, we see that it says there is no tiling that covers the mutilated board. This is precisely what we want the mutilated chessboard theorem to state.

# 3 THE PROOF

We now given a "semiformal" proof of our theorem couched strictly in terms of the Z formalisation. The semiformal proof comprises rigorous formal statements of a series of lemmas interleaved with narrative explanations of these statements and how they relate. The lemmas are given as Z conjecture paragraphs.

The informal proof uses the notion of the colour of a square that we have not yet formalised. The following integer-valued function on the squares represents this notion.

$$\begin{array}{|l} \boldsymbol{Colour} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} \\ \hline \forall x{:}\mathbb{Z} \times \mathbb{Z} \bullet Colour\ x = (x.1 + x.2)\ mod\ 2 \end{array}$$

This definition and elementary facts about remainders *modulo* 2 imply that the colour of any square is either 0 or 1:

$\boldsymbol{ran\_colour\_thm}$ $?\vdash \forall u{:}\mathbb{Z} \times \mathbb{Z} \bullet Colour\ u \in \{0,\ 1\}$

The following lemma gives a more explicit characterisation of adjacency saying that two squares are adjacent iff. their eastings differ by 1 and their northings are equal, or vice versa. It follows from the definition of adjacency and from elementary properties of the absolute value function.

$\boldsymbol{adjacent\_thm}$ $?\vdash$

$$\begin{array}{ll} \multicolumn{2}{l}{\forall x1,\ x2{:}\mathbb{Z} \times \mathbb{Z} \bullet} \\ & (x1,\ x2) \in Adjacent \\ \Leftrightarrow & abs(x1.1 - x2.1) = 1 \wedge x1.2 = x2.2 \\ \vee & abs(x1.2 - x2.2) = 1 \wedge x1.1 = x2.1 \end{array}$$

That every domino comprises two squares is wired into our definition of domino. That these two square have different colours needs to be proved. Using the definition of the colour function, case analysis given by *adjacent_thm* and elementary facts about remainders *modulo* 2, we have that the colours of any two adjacent squares are different:

$\boldsymbol{adjacent\_colour\_thm}$ $?\vdash \forall u,\ v{:}\mathbb{Z} \times \mathbb{Z} \mid (u,\ v) \in Adjacent \bullet Colour\ u \neq Colour\ v$

From the above facts and the definition of *Domino*, using elementary facts about sets of cardinality 2, we conclude that every domino comprises two squares, one with colour 0 and one with colour 1. This is the only fact about dominoes that will be used hereafter.

$\boldsymbol{domino\_colour\_thm}$ $?\vdash$

$$\forall x{:}Domino \bullet \exists u,\ v{:}x \bullet x = \{u,\ v\} \wedge Colour\ u = 0 \wedge Colour\ v = 1$$

If $z$ is a tiling then, within the squares it covers, there is a 1-1 correspondence between those coloured 0 and those coloured 1.

$\boldsymbol{tiling\_colour\_thm}$ $?\vdash$

$$\forall z{:}Tiling \bullet \exists f{:}\{u{:}\bigcup z \mid Colour\ u = 0\} \rightarrowtail\!\!\!\!\twoheadrightarrow \{u{:}\bigcup z | Colour\ u = 1\} \bullet true$$

The proof of the above is very simple and direct and does not require $z$ to be finite: using the definition of a tiling together with what we have proved about the colour function and elementary facts about bijections, it is routine to verify that we may take the following definition for a witness:

$$f = \{d{:}z;\ u,\ v{:}\mathbb{Z} \times \mathbb{Z} \mid d = \{u,\ v\} \land \mathit{Colour}\ u = \mathit{0} \land \mathit{Colour}\ v = \mathit{1} \bullet (u,\ v)\}$$

It follows that if the set of squares covered by a tiling is finite, then the colour function divides it into two subsets of equal size:

***tiling_colour_size_thm*** $?\vdash$

$$\forall z{:}\mathit{Tiling} \mid \bigcup z \in (\mathbb{F}\_) \bullet \#\{u{:}\bigcup z \mid \mathit{Colour}\ u = \mathit{0}\} = \#\{u{:}\bigcup z \mid \mathit{Colour}\ u = \mathit{1}\}$$

The proof of the above is immediate from elementary facts about finite cardinalities and the existence of the bijection between the two sets.

We now need to show that the mutilated board is not divided into two equal subsets by the colour function. To do this, we will first show that the unmutilated board can be so divided, which we can conveniently prove using *tiling_colour_size_thm*. To apply that lemma, we must show that the unmutilated board is finite:

***board_finite_thm*** $?\vdash \mathit{Board} \in (\mathbb{F}\_)$

...and that it can be tiled:

***board_tiling_thm*** $?\vdash \exists z{:}\mathit{Tiling} \bullet \bigcup z = \mathit{Board}$

For the witness in the above lemma, one can observe that $\{(i,\ \mathit{2} * j),\ (i,\ \mathit{2} * j + \mathit{1})\}$ is a domino and that two distinct dominoes of this form cannot overlap and so take:

$$z = \mathbb{P}\mathit{Board} \cap \{i,\ j : \mathbb{Z} \bullet \{(i,\ \mathit{2}{*}j),\ (i,\ \mathit{2}{*}j{+}\mathit{1})\}\}$$

Thus we may apply *tiling_colour_size_thm* and conclude that the unmutilated board is indeed divided into equal size subsets by the colour function:

***colours_on_board_thm*** $?\vdash$

$$\#\{u{:}\mathit{Board} \mid \mathit{Colour}\ u = \mathit{0}\} = \#\{u{:}\mathit{Board} \mid \mathit{Colour}\ u = \mathit{1}\}$$

The mutilated board is obtained by removing two squares of colour 0, and so we have:

***colours_on_mutilated_board_thm*** $?\vdash$

$$\begin{aligned} & \#\{u{:}\mathit{MutilatedBoard} | \mathit{Colour}\ u = \mathit{0}\} = \#\{u{:}\mathit{Board} \mid \mathit{Colour}\ u = \mathit{0}\} - \mathit{2} \\ \land\quad & \#\{u{:}\mathit{MutilatedBoard} | \mathit{Colour}\ u = \mathit{1}\} = \#\{u{:}\mathit{Board} \mid \mathit{Colour}\ u = \mathit{1}\} \end{aligned}$$

The mutilated chessboard theorem now follows, since *tiling_colour_size_thm* says that if a tiling of the mutilated board existed then the colour function would divide the mutilated board into two sets of equal size, whereas *colours_on_board_thm* and *colours_on_mutilated_board_thm* show this is not the case.

# 4 THE MECHANIZED PROOFS

The source text of this document is actually a script from which all the formal material can be extracted and checked using ProofPower[1]. The source text also includes a proof script giving the commands to prove all the conjectures mechanically. The proof script has not been

[1]See `http://www.lemma-one.com/ProofPower/index/` for more information on ProofPower.

included in the printed form of the document for brevity (it comprises some 11 pages of ML code).

Case studies like this are useful for pointing up areas in the mathematical library that need more supporting theory. As Paulson [7] points out, finite cardinalities are tricky to reason about. Fortunately, all the groundwork for this had already been done in ProofPower and the present argument only required a few extras. The supporting theory that had to be developed comprised a lemma about absolute values, 4 lemmas about remainders *modulo 2* and 6 lemmas about finite cardinalities, see appendix A for more details.

The proof of the mutilated chessboard theorem as given in section 3 and the proofs of the supporting theory discussed in appendix A comprise about 270 and 95 lines of proof commands respectively. The proof script also includes a few dozen lines of "boiler-plate", e.g., code to collect statistics on the numbers of inference steps. The proofs themselves are quite routine.

The specification was originally formulated to correspond very closely with the treatment given in [5] and inductive methods were used for the proofs as in [7]. This was done for comparative reasons with the expectation that the inductive approach was likely to be harder than what we present above. Including time taken on supporting theory, this approach took about 10 hours' work spread over three afternoons. The result is longer and more complicated than what we have presented above.

The main complications are that you have to prove that the tilings of interest are finite[2]. to justify proof by induction over them and then use induction to prove *tiling_colour_size_thm*. This requires more work in the proof itself and more supporting theory (e.g., to reason about the finiteness of finite powersets).

The proof was then reworked to use the argument given in section 3 above. This amounted to proving what we have called *tiling_colour_thm* and using it to replace the relatively long inductive proof of *tiling_colour_size_thm* with a simple derivation based on elementary properties of finite cardinalities. This took about 2 hours and reduces the amount of supporting theory required to what is presented above. A final hour or so was then spent simplifying and abbreviating the specification and adopting the proofs to accommodate.

The de Bruijn factor[3] for the mutilated chessboard theorem as presented here is, in my opinion, quite good. The same is true for the NQTHM, Mizar and Isabelle treatments. To estimate fairly, I feel that one must compare the formal treatment with a reasonably detailed exposition of the discrete mathematics along the lines of section 3 above. For the treatment in this document, the de Bruijn factor turns out to be roughly 3 measured either in terms of A4 pages (6.5:3) or in terms of lines of ML code versus lines of LaTeX source (270:110).

[2]Even closer to the approach of [7] would be to define the set of tilings as an intersection giving the smallest set of domino configurations containing the empty configuration and closed under the operation that adds a non-overlapping domino to a configuration. This saves having to prove that the resulting tilings are finite, but at the expense of making the specification more complicated and less intuitive and the results less general. It still results in a longer proof than the one presented above.

[3]This term, due to Freek Wiedijk, denotes the ratio between the size of a formal mathematical text and the size of its informal original.

# 5 CONCLUDING REMARKS

Further work, of course, remains to be done! Putting aside certain methodological issues in the formalisation, a significant shortfall is that both the informal and the formal accounts deal only with dominoes that lock into place over two adjacent squares of the board. I.e., both accounts tacitly make a discrete abstraction of what is intuitively a continuous problem. One must surely consider the possibility of a tiling of the board in which dominoes can be placed at arbitrary positions and orientations.

The more general problem can be reduced to the discrete one, e.g., using the observation that no bounded subset of the plane can be tiled in more than one way using like squares. Here, by "tiling", I mean simply a decomposition into squares whose interiors do not overlap. The tiling is not assumed to have any regularity and two squares may overlap along their boundaries in any way. To those whose systems support enough real analysis to consider it, I commend the geometric challenge of this reduction as food for thought.


## Acknowledgments

Thanks to Roger Bishop Jones and Larry Paulson for their comments.


# References


[1] R.D. Arthan and R.B. Jones. Z in HOL in ProofPower. *BCS FACS FACTS*, 2005-1.

[2] Gzregorz Bancerek. The Mutilated Chessboard Problem — checked by Mizar. In *[3]*. 1995.

[3] Robert Boyer and Andrzej Trybulec. *QED Workshop II*. Available on the World Wide Web at `http://www-unix.mcs.anl.gov/qed/`, 1995.

[4] International Standards Organisation. *Information Technology — Z Formal Specification Notation — Syntax, Type System and Semantics*. ISO/IEC 13568:2002.

[5] John McCarthy. The Mutilated Checkerboard in Set Theory. In *[3]*. 1995.

[6] John McCarthy. A Tough Nut for Proof Procedures. *Memo 16, Stanford Artifical Intelligence Project*, July 1964.

[7] Lawrence C. Paulson. A Simple Formalization and Proof for the Mutilated Chess Board. *Logic Journal of the IGPL*, 9(3):499–509, 2001.

[8] Piotr Rudnicki. *The Mutilated Checkerboard Theorem in the Lightweight Set Theory of Mizar*. Available on the World Wide Web at `http://www.cs.ualberta.ca/~piotr/Mizar/Mutcheck/`, 1995.

[9] J.M. Spivey. *The Z Notation: A Reference Manual, Second Edition*. Prentice-Hall, 1992.

[10] Sakthi Subramanian. *A Mechanically Checked Proof of the Mutilated Checkerboard Theorem*. Supplied with the NQTHM distribution, available on the World Wide Web at `http://www.cs.utexas.edu/users/boyer/ftp/nqthm/`, 1994.

# A SUPPORTING THEORY

In the narrative in section 3 above, we have appealed to elementary properties of the absolute value function, of remainders *modulo* 2 and of finite cardinalities. In this appendix we state the lemmas that were needed in the mechanised proof in addition to what is already provided in ProofPower.

## A.1 Absolute Values

The absolute value function is not defined in the standard Z toolkit but is available in the ProofPower Z library. In addition to the existing theorems about it, we need the following:

***z_abs_eq_1_thm*** ?⊢
$$\forall i,\ j{:}\mathbb{Z} \bullet abs\ (i - j) = 1 \Leftrightarrow i < j \land j = i + 1 \lor j < i \land i = j + 1$$

The proof is immediate by arithmetic after a case analysis on the sign on the operand allowing the absolute value function to be eliminated.

## A.2 Remainders *modulo* 2

Our first fact about remainders *modulo* 2 just specialises an existing theorem characterising remainders modulo $d$ for any non-zero $d$. The other results are easy consequences of this.

***z_mod_2_thm*** ?⊢ $\forall i,\ j{:}\mathbb{Z} \bullet i\ mod\ 2 = j \Leftrightarrow j \in \{0,\ 1\} \land (\exists d{:}\mathbb{U}\bullet\ i = 2{*}d + j)$

***z_mod_2_thm1*** ?⊢ $\forall i{:}\mathbb{Z} \bullet (2{*}i)\ mod\ 2 = 0 \land (2{*}i + 1)\ mod\ 2 = 1$

***z_ran_mod_2_thm*** ?⊢ $\forall i{:}\mathbb{Z} \bullet i\ mod\ 2 \in \{0,\ 1\}$

***z_plus_1_mod_2_thm*** ?⊢ $\forall i{:}\mathbb{Z} \bullet \neg i\ mod\ 2 = (i + 1)\ mod\ 2$

## A.3 Finite Cardinalities

The basics of the theory of finite cardinality are already provided, including the fundamental fact that if $a$ and $b$ are finite sets then $a \cup b$ and $a \cap b$ are also finite and one has $\#(a \cup b) + \#(a \cap b) = \#a + \#b$. From this material, we derive the following facts:

***z_size_pair_thm*** ?⊢ $\forall x,\ y{:}\mathbb{U} \mid \neg x = y \bullet \{x,\ y\} \in (\mathbb{F}\_) \land \#\{x,\ y\} = 2$

***z_size_eq_thm*** ?⊢ $\forall a,\ b{:}(\mathbb{F}\_) \mid a \subseteq b \land \#a = \#b \bullet a = b$

***z_size_2_thm*** ?⊢ $\forall a{:}(\mathbb{F}\_) \bullet \#a = 2 \Leftrightarrow (\exists x,\ y{:}\mathbb{U}\bullet\neg x = y \land a = \{x,\ y\})$

***z_finite_subset_thm*** ?⊢ $\forall a{:}(\mathbb{F}\_);\ b{:}\ \mathbb{U} \mid b \subseteq a \bullet b \in (\mathbb{F}\_)$

The following are the only theorems in this note that are proved by induction.

***z_size_product_thm*** ?⊢ $\forall a : (\mathbb{F}\_);\ b{:}(\mathbb{F}\_) \bullet a \times b \in (\mathbb{F}\_) \land \#(a \times b) = \#a * \#b$

***z_finite_size_thm1*** ?⊢ $\forall a{:}\ (\mathbb{F}\_) \bullet (\exists f{:}1\ ..\ \#a \rightarrowtail\!\!\!\!\rightarrow a\ \bullet true)$