INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9315498

Mathematical Go: An analysis

Rappaport, Melvin, Ph.D.

City University of New York, 1993

Copyright ©1993 by Rappaport, Melvin. All rights reserved.

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Mathematical Go: An Analysis

by

Melvin Rappaport

A dissertation submitted to the Graduate Faculty in Computer Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1993

COPYRIGHT 1993

Melvin Rappaport

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Computer Science in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1/25/93
Date

Michael Anshel
Chair of Examining Committee.

1/26/93
Date

Stanley Habib
Executive Officer

Dorian Goldfeld

Keith Harrow

Robert G. High
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Acknowledgements

This thesis is dedicated to Robert G. (Bob) High, whose untimely death in a rafting accident in the Chilean Andes brought great sorrow to people in many communities.

Bob had helped me with my research into go and Mathematical Go long before he became a member of my thesis committee. As a talented mathematician, delightful writer, strong go player, and dedicated populizer of Mathematical Go, he provided invaluable assistance on all aspects of my dissertation. It was Bob who gave me the impetus to write something that might be useful to go players.

I thank my mentor, Michael Anshel. He gave me direction when I was lost. He reignited my enthusiasm. He allowed me the freedom to experiment, yet kept me aware of the ultimate goal.

I thank Dorian Goldfeld. He invited me to sit in on his conversations with Elwyn R. Berlekamp on several occasions and arranged for Berlekamp to speak at Columbia, where I met other people who shared my interests.

I thank Keith Harrow. His tireless insistence on precision and clarity through multiple revisions was essential to my successful completion of this dissertation.

I acknowledge the kindness of Sammy Park whose patience and instruction, along with the support of other members of his go club, the Korea Baduk, took me from a total novice to an intermediate player.

I thank the members of the Brooklyn Go Club, whose camaraderie allowed me to test out a host of half-baked ideas.

I thank my wife Nancy.

Preface

And David slew Goliath. Thus ends a story familiar to many of us about two-person conflict. The setting was the Middle East, the time was several thousand years ago. David was a shepherd boy, bringing supplies to the army of his country. Goliath was the champion of an army at war with David's country. Goliath was a huge man with much armor for his body and head, and with a spear so large few men could carry it, let alone wield it effectively. David came equipped with a slingshot plus the knowledge and experience to use it.

And David slew Goliath. The traditional commentary says that David was armed with faith in his God. An alternative view is that David was also equipped with a knowledge of the probable outcome of the conflict, was equipped with a better strategic evaluation of the conflict, was in the possession of a superior tactical armament.

It would have been more peaceful if, after the battle of each country's champions, one side had acquiesced to the other's desires. That was not the case. After the individual combat, both nations conducted a bloody battle.

The history of mankind is replete with conflicts and battles that were settled in favor of superior technology. On the other hand, brute force occasionally triumphed. What made David's conquest memorable was the apparent inequity in the strength of both combatants at first appraisal.

And Berlekamp overcame the go professionals. Thus ends an as yet unfamiliar story about another triumph of technology over traditional experience. Elwyn R. Berlekamp is a famous mathematician known for, among other accomplishments, writing the definitive book on algebraic coding theory, and co-authoring with John Conway and Richard Guy the book "Winning Ways for your mathematical plays" [BE1] (henceforth referred to as "WW"). Berlekamp claims to be a weak go player [BE3]. Go professionals study for many years. They are able to memorize hundreds of complete go games. (An occidental learning the game may take several years to reach a level at which he can memorize just one game.) They spend thousands of hours studying endgame positions, memorizing patterns and shapes. Berlekamp recognized certain go endgame positions as extensions of WW methods and theories; he calls these positions "Mathematical Go." Some of these positions possess a depth of subtlety unknown to classical go analysis. Berlekamp was able to repeatedly beat go professionals taking either side of these special positions.

The technology needed to repeat Berlekamp's experiment does not require years to master. What Berlekamp discovered was a new view for an old situation. He saw that a go endgame situation can be isomorphic with WW's disjunctive sum of games. With moderate effort, the careful reader of this thesis with minimal go knowledge will be able to construct and play go endgame problems that will stump go competitors many times stronger than he.

This thesis looks at how to analyze Mathematical Go endgames, how to construct them, and how to play them. It discusses Mathematical Go's impact on the classical game of go and computer go. Chapter 1 defines a WW game, then discusses a typical game, Domineering. The intent is to provide an intuitive feel for a self-partitioning game that conforms to classical WW rules. Chapter 2 presents the relevant WW theory of games needed for Mathematical Go. Chapter 3 describes the rules of go and Mathematical Go. Chapter 4 analyzes Berlekamp's only published example of Mathematical Go. Chapter 5 reviews some of David S. Wolfe's doctoral thesis on Mathematical Go. Chapter 6 relates Mathematical Go to the classical game of go. Chapter 7 discusses how to construct Mathematical Go problems. Chapter 8 discusses a special go rule, ko, and its limiting effect on extending Mathematical Go. Chapter 9 discusses how some WW ideas can be applied to a broader range of go positions than pure Mathematical Go. Chapter 10 discusses Mathematical Go's possible impact on computer go.

TABLE OF CONTENTS

Acknowledgements iv

Preface vi

Table of Contents ix

List of Figures x

Chapter 1. Classical Winning Ways theory and a typical game - Domineering.............................. 1

Chapter 2. Games, trees, values, and notation - more Winning Ways theory 12

Chapter 3. Go and Mathematical Go 25

Chapter 4. A Mathematical Go Problem and its solution 55

Chapter 5. Corridors and sockets 81

Chapter 6. MathematicalGo values and go scoring 90

Chapter 7. Design Issues in Mathematical Go 103

Chapter 8. Ko 119

Chapter 9. Sente and Gote 127

Chapter 10. Computer Go 133

Epilogue 141

Appendix 145

Bibliography 148

List of Figures

Figure 1.1 Domineering 4

Figure 1.2 An example of a self_partitioning game 6

Figure 1.3 The four youngest WW games 9

Figure 1.4 Not all games partition themselves 11

Figure 2.1 Games of length 2 13

Figure 2.2 Bypassing a reversible move 16

Figure 3.1 A corner, side and middle play 28

Figure 3.2 Captures and kos 32

Figure 3.3 Two eyes central, mimimal stones 36

Figure 3.4 Two eyes central, encircling stones 36

Figure 3.5 An unusual living group 38

Figure 3.6 Live corner and side groups 39

Figure 3.7 Two, three, and four liberties in a row 40

Figure 3.8 Nakade 42

Figure 3.9 Seki... 45

Figure 3.10 Endless Life 47

Figure 3.11 Trople ko.................................... 48

Figure 3.12 Suicide 51

Figure 4.1 Berlekamp's problem 56

Figure 4.2 The problem partitioned into tiles 59

Figure 4.3 Settled territory 62

Figure 4.4 Tile C and Tile D 64

Figure 4.5 Tile E 65

Figure 4.6 Tile B 67

List of Figures (continued)

Figure 4.7 Tile F 68

Figure 4.8 Tile A unsimplified 69

Figure 4.9 Tile A simplified 70

Figure 4.10 Game summary 77-79

Figure 5.1 The chilled value of corridors 84

Figure 5.2 Group invading many corridors 87

Figure 5.3 Multiple sockets 88

Figure 6.1 A chilled corridor of length 3 93

Figure 6.2 Playing the sum of 4 tiles 96

Figure 7.1 Symmetry....................................... 110

Figure 7.2 The T2 tiles 111

Figure 7.3 The T3 tiles................................... 113

Figure 7.4 The 5 topological shapes of size 4 114

Figure 8.1 The T2.2 tile 120

Figure 8.2 A game fragment 120

Figure 8.3 The sum of three kos 122

Figure 9.1 A position by Sakata 130

Figure A.1 A position by Takeo 146

Figure A.2 A position by Kensague 146

CHAPTER 1. Classical Winning Ways theory and a typical game - Domineering

Let us begin by examining the roots of Mathematical Go. Mathematical Go is a two-person game of perfect information. The players, Black and White, each have different moves available to them, and each knows the other's possible replies. Mathematical Go is based on the theories developed in WW [BE1, Volume 1, Chapters 1-6]. WW made a significant extension of the number of games analyzed in detail, especially games which are not impartial.

An impartial game is defined as a game with the same moves available to both players at their respective turn to play. In the 1930's R.P. Sprague and P.M. Grundy showed that every impartial finite game for two players reduces to a Nim-like game. To use the terminology of WW [BE1, p.58]:

> Every impartial game is just a bogus Nim-heap (that is, a Nim-heap with reversible moves added from some positions).

The classical game of Nim (there are many variants) has several heaps of objects. Players alternate moves by removing some or all of the objects from any one heap. The first player who cannot make a move (because no objects are left) loses. Bogus Nim-heaps arise when a player is allowed to add objects to a heap as long as the rules imply some "minimally excluded number" of objects that he is not allowed to create [BE1, p.57].

The value of any Nim heap is what WW calls a "nimber" [BE1, p.43]. To obtain the sum of a collection of nimbers "write the nimbers in binary and add without carrying" [BE1, p. 75]. For example, the game with heaps of size 3, 4 and 5 evaluates to 11 + 100 + 101 = 212. A position with an odd total in a column is a win for the player on move, for it can always be transformed to a nimber consisting of all even columns. In the example above, removing 2 objects from the heap of size 3 evaluates to 1 + 100 + 101 = 202.

By knowing how to play Nim and having a mapping of positions in an impartial game to their Nim-heap value, one can play strategically perfect moves for that game. However the actual playing may involve considerable effort. As pointed out by Richard J. Nowakowski in "..., Welter's Game, Sylver Coinage, Dots-and-Boxes, ...," [NR, p. 155]

> Welter's Game is completely within the scope of the Sprague-Grundy theory, but to find the nim-value of a position and a good reply is quite difficult.

When the moves available are not the same for each player, then a game is not impartial. WW calls such a game "partizan" [BE1, p. 17]. The foundation of Mathematical Go is recognizing the values of certain partizan games and knowing how to sum a collection of such values. As we shall see, the possible values of partizan games are far more numerous than nimbers, and the rules for summing them are much more complex.

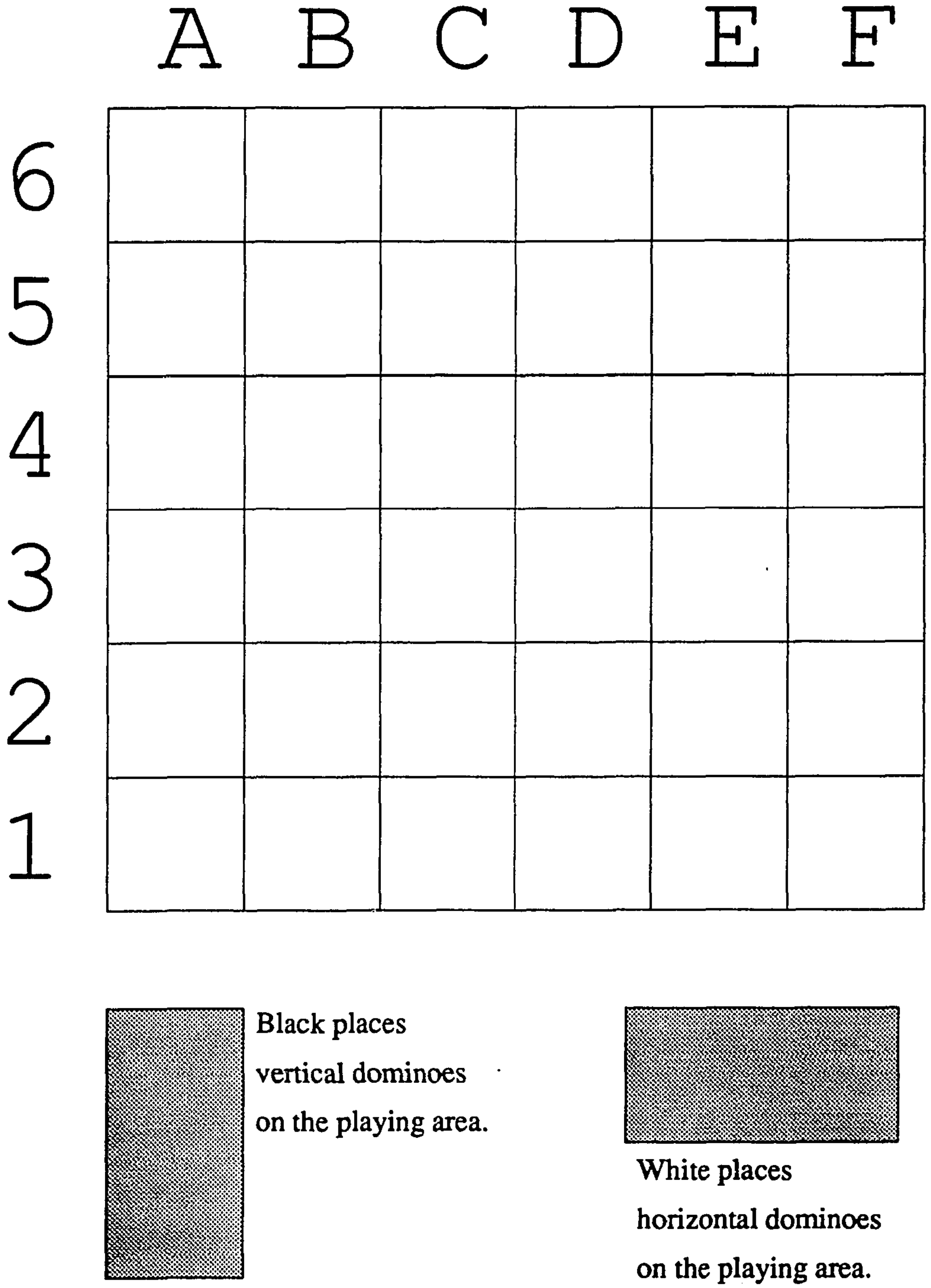
The traditional game of go is a partizan game. Black places black stones and White places white stones on a ruled board. Legal moves for each side are different and depend on the patterns of the previously played stones. Mathematical Go, an approximate subset of go, is a partizan game. Checkers and chess are partizan games for similar reasons.

A game that always ends is a finite game. Nim is finite. Go, Mathematical Go, chess, and checkers are not finite because a potential draw by endless repetition of moves is possible. WW's approach for near-optimal strategy in playing games was restricted to finite games. Berlekamp's Mathematical Go represents a significant extension to WW theory, presenting an optimal winning strategy for positions whose variations include possible draws.

Until this point, we have been talking about games without providing a definition. It is now time to look at the definition of a game as presented in WW. In a WW game [BE1, p.16]:

1. There are just two players, often called Black and White.
2. There are several, usually finitely many, *positions*, and often a particular *starting position.*
3. There are clearly defined *rules* which specify the *moves* that either player can make from a given position to its *options.*
4. Black and White move alternately, in the game as a whole.
5. In the *normal play* convention a player unable to move *loses.*
6. The rules are such that play will always come to an end because some player will be unable to move. This is called the *ending condition.* So there can be no games which are drawn by repetition of moves.
7. Both players know what is going on, i.e. there is *complete information.*
8. There are no *chance moves* such as rolling dice or shuffling cards.

To see WW strategy in action, lets examine a typical WW game. Domineering [BE1, p.117 ff.] is a game that is extremely easy to learn and illustrates many of the basic ideas used in later analysis. Domineering is played on a rectangular board, ruled in squares, like the familiar checkerboard. Black playing first can place a domino (a 1 x 2 rectangle) vertically on any empty location, White playing second can place a domino horizontally on any remaining location. No two dominoes may overlap. The game is lost by the player on play who has no legal move. Figure 1.1 illustrates the game board and the player's dominoes.


A
B
C
D
E
F
6
5
4
3
2
1

Black places vertical dominoes on the playing area.

White places horizontal dominoes on the playing area.

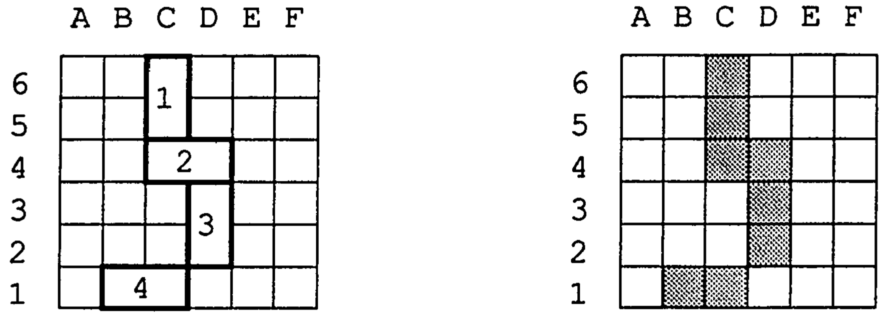
Figure 1.1 - Domineering

Figure 1.2 shows the start of a sample game. (No comment will be made on the merits of the play from a strategic viewpoint.) Numbers inside the board are used to number the plays. Coordinates of the board are supplied outside the playing field, should the reader wish to record moves; they are not used in this discussion. Black makes the odd-numbered moves 1,3, and 5; White makes the even-numbered moves 2, 4, and 6.

The game board before any moves has one region. After four moves this game as a whole has two mutually distinct regions, which we label "tiles," each subject to all of the rules of the game as a whole. After two more moves the board is now partitioned into four tiles. Again, each tile is subject to all of the rules of the game as a whole.

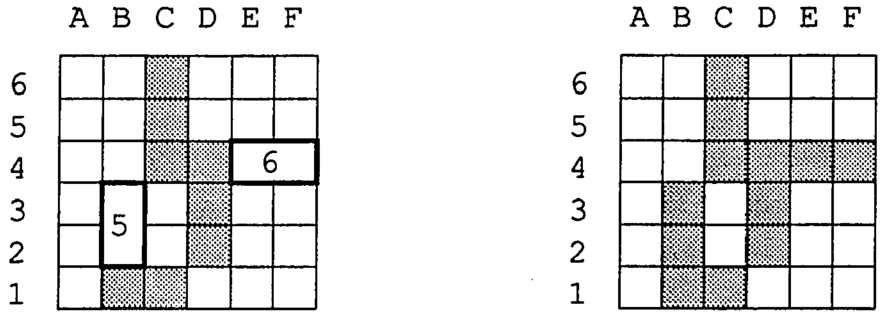
We use the term tile to mean a partitioned region of the playing area such that a move in one tile does not affect the legality of moving in any other tile. The use of the word "tile" emphasizes the particular morphology of a game. As we shall soon discover, different tiles can have the same value. WW uses the terms game, sub-game, position, component, follower, and option at various points to mean nearly the same thing as tile.

The result of a play in one tile will be one or several (two, three, or four) new tiles, but *tiles can never be combined.* For the purposes of analysis, one can record the play of a game by the changes in the configurations of its tiles. Draw a tree with the initial tile as the root. For each move in a tile draw an edge downwards from that tile to its newly formed tile or tiles. One can extend this tree from a game position to all future potential plays. Eventually the legal moves available from any original tile must be exhausted. This tree is one representation of a game tree. By WW convention, a Black move from a tile slopes to the left while a White move from a tile slopes to the right.


A B C D E F
6
5
4
3
2
1
1
2
3
4

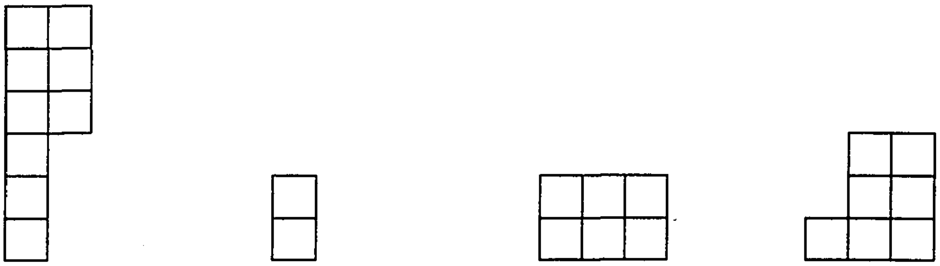
A- Four moves played

B- The game is partitioned into two tiles


A B C D E F
6
5
4
3
2
1
6
5

C- Two more moves played

D- The game is now partitioned into four tiles



Tile A

Tile B

Tile C

Tile D

A 6 x 6 Domineering board.

An example of a self-partitioning game

Figure 1.2

WW game trees for partitioned tiles are not the search trees used in minimax analyses. Search trees alternate layers of Black and White. Search trees always consider all moves everywhere from any position as part of one game. Thus, the fanout of an exhaustive search tree node depends on the total number of moves from anywhere in the playing area. Typically, this fanout is nearly constant to the very end of the game. Exact evaluation is often intractable. After a WW tile is partitioned into several tiles, the node for each child tile no longer has available the individual moves of its siblings, its fanout is significantly reduced. As tiles grow smaller, exact evaluation becomes feasible, then easy.

We have been describing a top-down approach to exhaustively enumerate all moves in Domineering (or any finite partizan game). Games of this type can also be analyzed from the bottom up. For each tile there exists an associated game tree. One can characterize a game tree by the length of its longest sequence of moves. WW calls game trees with shorter length "younger" than game trees of longer length. One first evaluates the youngest game trees, tiles where only one move is available. Next, one calculates values of older trees having more moves to completion and more initial choices, using the values already calculated for the younger trees. In Chapter 7, we shall use this bottom up approach, also known as dynamic programming, to outline a method for evaluating tiles occurring in Mathematical Go.

The value of a tree is related to the excess of moves a player has over his opponent, positive for Black and negative for White. Based on the value of the actual tiles in play, treating each tile as the root of a tree, one selects which tile to play in, then makes the best move in that tile. The best move in a tile is that which leaves the most favorable position for the player after a move has been made.

Figure 1.3 illustrates how values are assigned to some simple tiles. On the left are shown all tiles whose associated game tree has at most one move. On the right is the associated game tree and the WW set notation for that tree. This notation will be discussed in Chapter 2. The tile 1x1 offers 0 moves to either side and doesn't affect the rest of play; its value is 0. The vertical 2x1 tile with one move for Black and none for White has value 1. By convention, values favorable to Black are positive and values favorable to White are negative. The vertical 2Nx1 tile has value N; e.g., the tile with 10 squares arranged in one column has value 5. Similarly the horizontal 1x2 tile has value -1 and the horizontal 1x2N tile has value -N. Notice that there may exist several board representations that reduce to the same WW value. The three tiles with the value * are the simplest example of a tree value that cannot be represented as a number.

These four games represent the set of all possible outcomes for a finite game. These outcomes are:

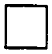
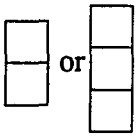
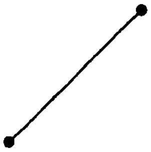
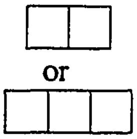
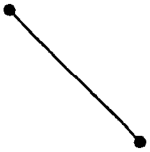
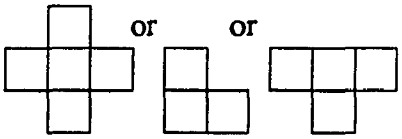
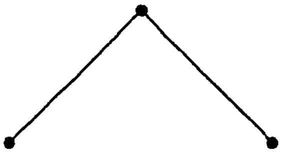
a win for Black,

a win for White,

a win for the first player,

a win for the second player.

Every game where the second player wins has a value of 0. Every game where Black wins has a positive value. Every game where White wins has a negative value. Every game where the first player wins has a value that when compared with zero is neither less than nor greater than 0. Games that can't be compared with 0, such as *, are called "fuzzy." Chapter 2 will discuss comparison of games.

<table>
<tr><th>Domineering tiles</th><th>Winning Ways Game Trees</th></tr>
<tr><td>
0 : pronounced zero,
a second player win</td><td>
{ | }</td></tr>
<tr><td>
1 : pronounced one,
a win for Black</td><td>
{ 0 | }</td></tr>
<tr><td>
-1 : pronounced minus one
or negative one,
a win for White</td><td>
{ | 0 }</td></tr>
<tr><td>
* : pronounced "star,"
A first player win</td><td>
{ 0 | 0 }</td></tr>
</table>

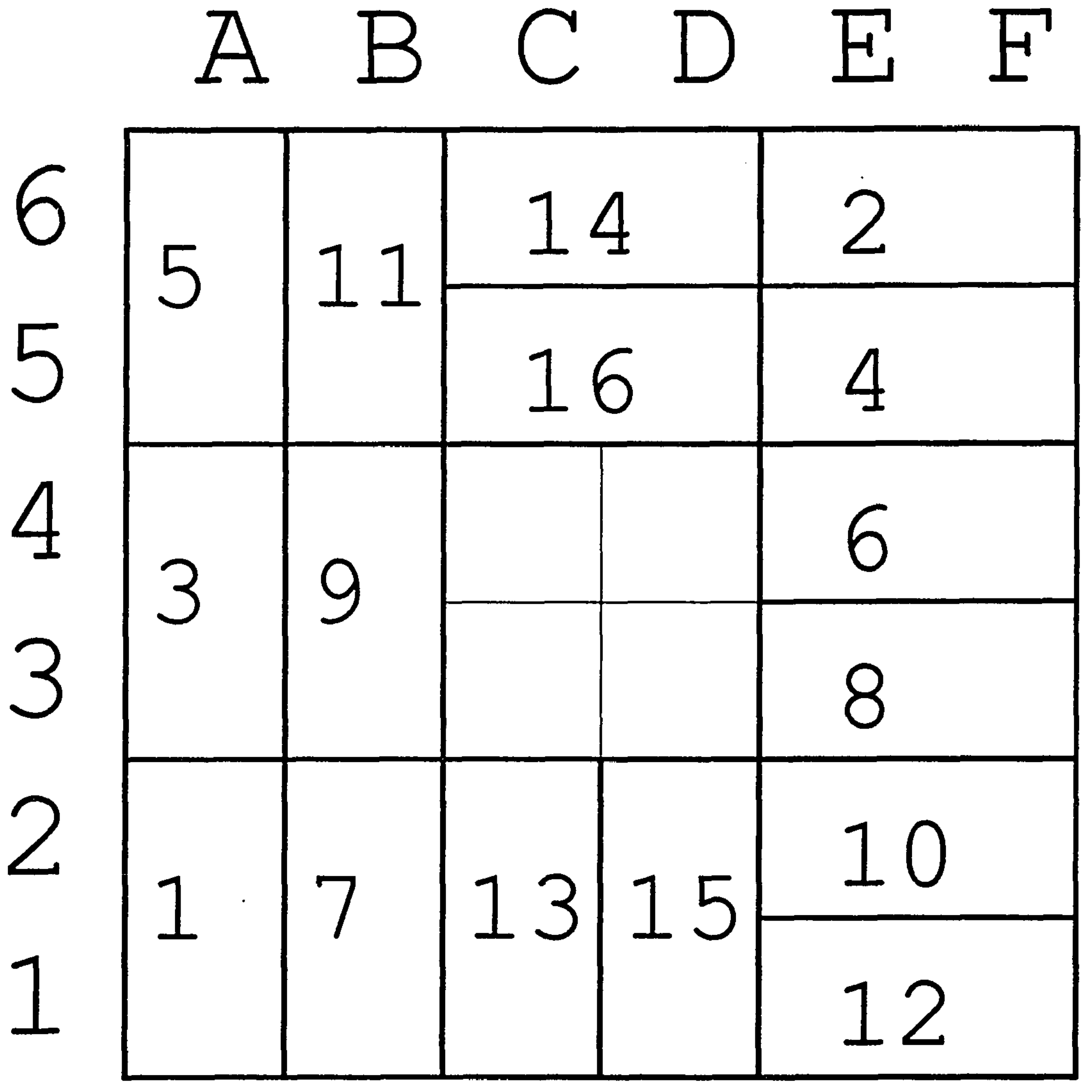
The four youngest WW games, illustrating all outcome classes
Figure 1.3

Using the familiar operation of addition, we can sum values of several tiles to arrive at a value for the compound game. A game whose tiles have values +3, +5, and -2 will have value of +6, which means Black will win with six moves to spare.

The primary lesson illustrated by Domineering is that the rules of the game apply to any isolated components ("tiles") which may emerge in the course of play as well as to the original position. The more a game can be partitioned into tiles, the more powerful will be the advantage to a player skilled in WW strategies, knowing the value of the separate tiles. Figure 1.4 is a sample game (again, no comment on the merits of the play) where no partitions occur. At no point in this game will the power of WW strategy be employed effectively, to select among separate tiles the tile that will furnish the best play. Of course, when a game is not partitioned during play, the methods of WW can be used to advantage when evaluating possible future moves that eventually lead to partitions, even if they don't get played.

To summarize this chapter, we have defined the rules of a WW game; discussed a typical game, Domineering; introduced a few young games with their values and outcome classes; shown top-down and bottom-up approaches to evaluating game trees (tiles); and anticipated the strategy of Mathematical Go by appealing to the reader's intuition that correctly evaluating a tile and knowing how to sum tiles allows for an exact strategy in small games.

The next chapter will place many of these ideas in a more rigorous setting.


A
B
C
D
E
F
6
5
4
3
2
1
5
11
14
2
16
4
3
9
6
8
1
7
13
15
10
12

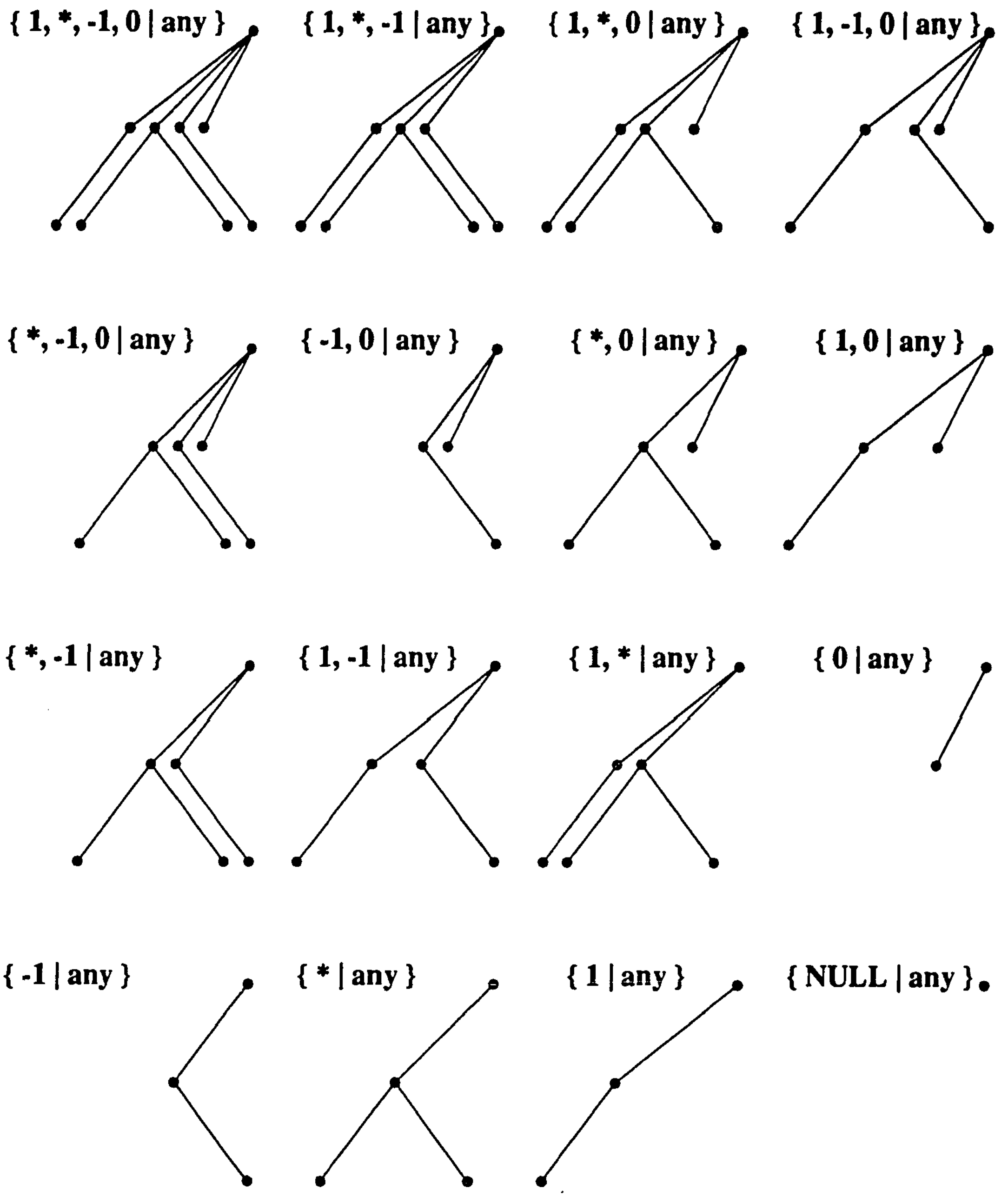
Not all domineering games partition themselves into tiles.
Figure 1.4

CHAPTER 2. Games, trees, values, and notation - more Winning Ways theory

Some basic results of WW are that every game has a value, values are additive, values have negatives, values can be simplified, and values are comparable. These concepts apply to all games, but specifically they form the basis for understanding Mathematical Go. This chapter discusses game values and the operations that are performed on game values.

In Chapter 1, we saw that some values were numbers and at least one game value, *, was "fuzzy," not comparable with zero. Indeed, it's value is not a number. In Chapter 1, we discussed the techniques for drawing game trees from the legal moves in a tile. We gave some examples of WW graphical trees and their associated set notation. Game trees provide a visual description of the value of a game. In WW, the name of the game tree and its evaluation are often used synonymously. A game tree is an abstraction of a tile, it concerns itself only with the number of moves left, and does not care about the physical configuration of the board. WW games can be played not only on boards, but with coins, cards, matchsticks, paper and pencil, and so forth. The pertinent abstraction will always be the game tree of moves remaining based on particular plays.

One can devise an algorithm for enumerating the game trees for all finite games. We have seen in Chapter 1 that there are four games with length less than or equal to 1, namely: 1, *, -1, and 0. In games of length less than or equal to 2, the player on move must move to one of these four games. However, depending on the set of rules for a particular game, a player might have the choice to move to some, all, or none of these games. Figure 2.1 illustrates this for the Black player. With the initial position as the root of the tree, there are sixteen distinct groupings of moves available, corresponding to the cardinality of the power set of a set with


{ 1, *, -1, 0 | any }
{ 1, *, -1 | any }
{ 1, *, 0 | any }
{ 1, -1, 0 | any }
{ *, -1, 0 | any }
{ -1, 0 | any }
{ *, 0 | any }
{ 1, 0 | any }
{ *, -1 | any }
{ 1, -1 | any }
{ 1, * | any }
{ 0 | any }
{ -1 | any }
{ * | any }
{ 1 | any }
{ NULL | any }

The 16 unsimplified choices for Black in games of length ≥ 2.
Also known as the power set for { 1, *, -1, 0 | }.
Figure 2.1

four elements. By symmetry, White will also have 16 grouping of moves available for every one of Black's 16 choices. Thus there are 16 x 16 = 256 distinct game trees with length less than or equal to two.

It's easy for most people to visualize the moves in a game from its game tree. However, to discuss arbitrary games, WW uses a set notation that closely maps to the game tree. This notation is used alongside the trees in Figure 2.1.

A WW game is recursively defined as [BE1, p.37]:

$$G = \{ G^B \mid G^W \}$$

This is to be read as: the game G is defined by an ordered pair of sets of games; G^B represents the possible resulting games after a legal move in G by Black, and G^W represents the possible resulting games after a legal move in G by White. After a player makes a move in G, the name/definition of the resulting game is the appropriate one game, either from G^B or from G^W, depending on which player moved. That one game is defined by its possible resulting games after a legal move by a player, just as all games are defined.

The games in G^B and G^W are represented in the same way. We saw in Figure 2.1 that the set of Black moves can be enumerated and separated by commas, and likewise for White. The enclosing braces are often omitted when no confusion might arise. In order to represent nested moves, that is sequences of moves longer than one, one may use paired braces or multiples of "|," the player delineator, one for each level of move. The fully expanded, non-recursive, set representation of a game is quite similar to the exhaustive game tree for that game, with each final element representing a leaf node of the game tree.

Figure 2.2 illustrates a game of length 3. We will discuss the meaning of the figure shortly. The set notation for the game G is:

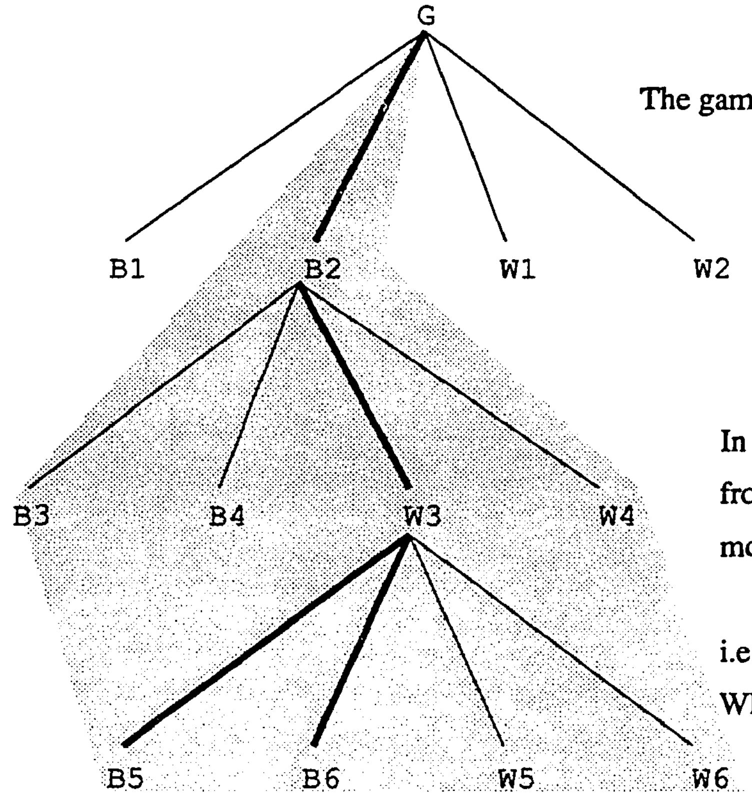
G = { B1, B2 | W1, W2 }
= { B1, { B3, B4 | W3, W4} | W1, W2 }
= { B1, { B3, B4 | { B5, B6 | W5, W6 }, W4} | W1, W2 }

In this case, removal of the braces would lead to confusion. However, the game 4|3|2|0 can uniquely be read as follows: as a first move, Black may move from this game to the game 4, White may move from this game to the game 3|2|0; again, after White's first move Black may move to the game 3 and White may move to the game 2|0; and so forth.

Best play assumption

There are 256 unsimplified games of length two, games that have distinct sets of moves available. However if one played these games with best play on both sides, only 22 distinct games will be played. This difference in number is explained by the best play assumption. For the purpose of evaluating a game, WW assumes that each player will always move to maximize his outcome, will never make a less than optimal move. Since all WW games involve perfect information, (i.e., there is never an outcome based on probability as in the games discussed in "Theory of Games and Economic Behavior" [VJ] by J. Von Neumann and O. Morgenstern), one can determine a certain move to be less than best play and discard it from consideration. For example in the game { 1, -1 | 0} Black may move to either 1 or -1, to either a game that Black always wins or a game Black never wins. As a reasonable player, Black selects the strategy to always move to 1. The game simplifies to the game { 1 | 0}.

The game G before simplification.

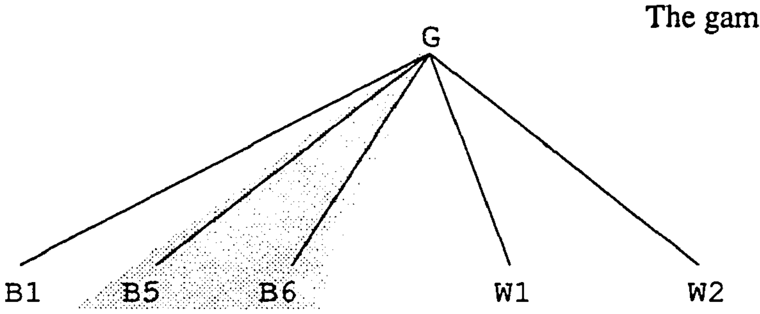

G
B1
B2
W1
W2
B3
B4
W3
W4
B5
B6
W5
W6

In order to reverse Black's move from G to B2, White must have a move from B2 to W3 such that:

$$G \geq W3$$

i.e., W3 is at least as good for White as was G, if not better.

The game G after simplification.


G
B1
B5
B6
W1
W2

The bottom game has exactly the same value as the top game.

Simplification: Bypassing a reversible move

Figure 2.2

Sums, negatives, and ordering

The basic operations on games are sum, negative, and comparison. The definition of the sum of two games is: [BE1, p. 33]

$$G + H = \{ G^B + H ,\ G + H^B \mid G^W + H ,\ G + H^W \}$$

This says that the sum of games G and H (or the effect of playing the two positions G and H as one combined game) is equivalent to the game where:

either Black will make a move in one game (either G or H),

replacing that game with a new position (a Black follower of that game)

while leaving the other game unchanged;

or White will make a move in one game (either G or H),

replacing that game with a new position (a White follower of that game)

while leaving the other game unchanged.

It follows from the definition that sums of games are commutative, associative, and have a zero, the game with no moves, $\{ \mid \}$, such that G + 0 = G.

Every game has a negative, which is defined as reversing the roles of the two players. Whatever moves that were previously legal for White are now legal for Black, and whatever moves that were previously legal for Black are now legal for White, that is:

$$-G = \{ -G^W \mid -G^B \}$$

It follows that G + (-G) = 0. This is easy to see strategically. Wherever the first player plays, the second player mimics that move in the corresponding other tile. The second player will always have a move since the initial positions were identical. If one turns to Figure 1.3 and then rotates the page by 90^o the negatives of each tile are displayed. It is easy to verify that:

$-(0) = 0$, $-(1) = -1$, $-(-1) = 1$, and $-(*) = *$.

With best play on both sides, every game G is a member of exactly one of four possible outcome classes:

zero - if the second player wins ($G = 0$).

positive - if Black wins ($G>0$).

negative - if White wins($G<0$).

fuzzy - if the first player wins(G 0, also pronounced gee confused with zero).

These classes may be combined as:

$G \geq 0$ means $G > 0$ or $G = 0$.

$G \leq 0$ means $G < 0$ or $G = 0$.

$G \,\|\!> 0$ means $G > 0$ or $G \parallel 0$.

$G <\!\| \, 0$ means $G < 0$ or $G \parallel 0$.

The definition of comparison of games is a constructive one.

$G = H$ is verified by $G + (-H) = 0$.

$G > H$ is verified by $G + (-H) > 0$.

$G < H$ is verified by $G + (-H) < 0$.

$G \parallel H$ is verified by $G + (-H) \parallel 0$.

$G \geq H$ is verified by $G + (-H) \geq 0$.

$G \leq H$ is verified by $G + (-H) \leq 0$.

$G \,\|\!> H$ is verified by $G + (-H) \,\|\!> 0$.

$G <\!\| \, H$ is verified by $G + (-H) <\!\| \, 0$.

Constructive comparison means to verify a relation between games, play out the games as indicated, then check the outcome class. Thus to show G > H, find the negative of H, namely (-H), then verify the outcome of the sum of games G + (-H) is always a win for Black. This verification may involve exhaustive enumeration of the sum of games, or may use simplification (see below). The other relations are verified by the same construction.

Among the first four games discussed in Figure 1.3 there is a partial ordering:

1 > 0, 1 > *, 1 >- 1, 0 > -1, * > -1, and *‖0.

These are all readily verified by playing in the appropriate sum of games. Reviewing the 16 distinct games of Figure 2.1, Black would always move to the most positive game, so only six distinct "best play game trees" remain; namely:

{1 | any}, {0 | any}, {* | any}, {0,* | any}, {-1 | any}, and {NULL | any}.

These six choices are also available to White.

The reduction of exhaustive game trees to best play game trees is called "simplification." There are two methods for simplifying a game, by deleting dominated options, and by bypassing reversible moves.

If G = {A,B,C,...|D,E,F,...} and A ≤ B and D ≥ E, then the value of the game H = {B,C,...|E,F,...} = G. The option B dominates A, and the option E dominates D. The game H has deleted the dominated options and maintained the same value as G.

Figure 2.2 illustrates the other method for simplification. If Black has a move such that White has a reply leaving a position as least as good for White as the original position then Black's move is reversible. In Figure 2.2, assuming G ≥ W3, then Black's move from G to

B2 is termed "reversible." White's reply from B2 to W3 is the reason for the reversibility, and the game illustrated in the lower half of Figure 2.2 has the same value as the original game. The change in the game tree is called "bypassing a reversible move." The same procedure applies with the colors reversed.

From a player's viewpoint, the bypassed Black move may be at any level of tactical or strategic strength: weak, temporizing, or strong. The key to this sequence is that White's reply makes White's position no worse than it was before the two move sequence. Since White improves (or maintains) his value, we can consider his reply as made immediately.

Note that in Figure 2.2, the morphology of the game tree changes after simplification, even though the value is preserved. In the simplified tree the moves to B2 and W3 are lost. When game trees for Mathematical Go are discussed, we shall adopt the convention of:

> preserving bypassed nodes such as B2 and W3, and preserving the paths through them as thickened edges, such as G to B2 and B2 to W3. The edges to B3, B4, W4, W5, and W6 along with those nodes will either be dashed lines or omitted. Similarly deleted dominated moves will be dashed lines or omitted.

Thus one may follow the play through a reversible sequence from the game tree.

Conway, in "On Numbers and Games" proves several results about simplification:

1) Deleting dominated moves and bypassing reversible moves from a game does not change the value of a game, it only removes some choices in play.

2) Games with the same value after simplification are identical.

3) Games simplified by deleting dominated moves and bypassing reversible moves are as simple as they can be.

We had shown six distinct games for Black of length 2 after simplification by deleting dominated moves. Multiplying by White's six games yields 36 games. Simplifying our 36 games using the methods of bypassing reversible moves, only 22 distinct games remain. These are:

{1|1}, {1| }, {1|*}, {1|0}, {1|0,*}, {1|-1},
{0|1}, {0| }, {0|*}, {0|0}, {0|0,*}, {0|-1},
{0,*|0,*}, {0,*|0}, {0,*|-1}, {*|0}, {*|-1},
{-1|0}, {-1|-1}, { |0}, { |-1}, { | }.

Using this result, it is easy to estimate the number of unsimplified game trees of length 3. The cardinality of the power set for a set with 22 elements is 2^{22}. Since either Black or White may have this choice, then there are 2^{22} x 2^{22} or approximately 16,000,000,000,000 unsimplified games. Berlekamp [BE3] has stated that the number of simplified games of length 3 is less than 1,000. Using his estimate, the number of unsimplified games of length 4 is on the order of 2^{1000} x 2^{1000} or 10^{600}. The point of these estimations is that the total number of distinct finite partizan games, whether simplified or unsimplified, grows at a hyper-exponential rate with the number of moves. However, a specific game, such as Domineering, will not have such a rapid growth in the number of game tiles with distinct values because the rules of the game will not allow the creation of most WW game trees. Thus, enumeration of tiles may be a feasible approach to the analysis of certain games. This will be discussed for Mathematical Go in Chapter 7.

Canonical forms, numbers, hot and cold games

By repeatedly applying the two methods of simplification, in any order, at any part of the game tree, a point will be reached where no further simplification can be made. The resulting simplest game is known as the canonical form for the original game. It is unique and corresponds to the WW tree of shortest length with value equal to that of the original game.

We have seen that the integers occur as values of certain games. WW shows that fractions also appear as the values of certain games. Using canonical forms of finite games one can express: 0, all positive and negative integers, and all positive and negative dyadic fractions (simple fractions with the denominator a power of two). One can generate all these numbers with the following rules [BE1, p. 24]:

$$0 = \{ \mid \}$$

$$n+1 = \{ n \mid \}$$

$$-n-1 = \{ \mid -n \}$$

$$n + 1/2 = \{ n \mid n+1 \}$$

$$\frac{2p+1}{2^{n+1}} = \{ \frac{2p}{2^{n+1}} \mid \frac{2p+2}{2^{n+1}} \} = \{ \frac{p}{2^n} \mid \frac{p+1}{2^n} \}$$

We have defined the "simplest" form of a game as the fully simplified game. The simplest form of a game has the exact same value as the game before reduction, and when played out, has the fewest possible moves. When both G^B and G^W are integers or dyadic fractions, the value of G is the simplest game between the two values. For example,
$\{1/4 \mid 1\} = 1/2$, $\{7/4 \mid 6\} = 2$, $\{-40 \mid +100\} = 0$.

For those who didn't easily grasp the above examples, let's list the values for short games whose values are numbers. For convenience, only positive numbers are shown, their negative values have the same game tree length. To follow the convention of "On Numbers and Games," [CJ] we label games with tree length *n* as day *n*.

day 0: game 0.

day 1: game 1.

day 2: games $\frac{1}{2}$, 2

day 3: games $\frac{1}{4}$, $\frac{3}{4}$, 1 $\frac{1}{2}$, 3

Look for numbers that fit between G^B and G^W, starting with the earliest day. In the first example, O doesn't fit, 1 doesn't fit, $\frac{1}{2}$ does fit, so it is the correct value. In the first example, O doesn't fit, 1 doesn't fit, 2 does fit, so it is the correct value. In the third example, a bit tricky, we have a second player win, or 0, which fits. The operation of evaluating games is *not* just the average of the two numbers.

The Number Avoidance Theorem

Any game G, whose canonical form has a numeric value, is of the form G^B < G < G^W. Any Black move will lead to a game less favorable to Black, any White move will lead to a game less favorable to White. A game with a value that is a number is called a "cold game." A game that is not a cold game is called a "hot game."

G = { 0 | 1 } is a cold game with value 1/2. When either player makes a move in G he suffers a net loss of 1/2. H = { 1 | 0 } is a hot game with no numeric value (it does have a "mean value" of 1/2). When either player makes a move in H he profits by roughly 1/2. The

Number Avoidance Theorem [BE1, p. 144] tells us when playing in a sum of hot and cold games, don't play in a cold game unless there's nothing else to do. This applies in all cases.

The concepts developed in this chapter are used in analyzing Mathematical Go in Chapters 4, 5 and 6. Additional material from WW will be presented within the specific context of Mathematical Go. Before the analysis, Chapter 3 will present the rules of go and Mathematical Go.

CHAPTER 3. Go and Mathematical Go

Prior to Berlekamp's investigations into Mathematical Go, the most popular games analyzed using WW theory were Nim and dots-and-boxes. These games are not widely played at clubs, there are no major tournaments, they are not discussed in newspapers and magazines. Indeed, of all the many games discussed in WW, not one would be called a major recreational pursuit among game players of the world.

Go is played throughout the world by tens of millions of people. Large tournaments have prizes worth hundreds of thousands of dollars. In the Orient, many newspapers have regular columns on go, and go is featured on television and in magazines. In 1991 Kobayashi Koichi, a Japanese go professional, earned $785,000 in tournament prizes. This does not include endorsements for personal appearances, speaking fees, and other accoutrements of success in a major sport. Ing Chang-Ki, a successful Taiwanese businessman, has offered a prize now worth about $1,600,000 should a computer go program reach professional strength. Berlekamp's new outlook on the ancient game of go has greatly enlarged the potential audience for WW theory.

The rules of Mathematical Go are very similar to the rules of go. This chapter will first present the rules of go, emphasizing those features that are discussed in subsequent chapters. Next, the rules of Mathematical Go are presented.

Go is played throughout the world with slight variations in the rules. Berlekamp calls these different sets of rules "dialects" [BE2, p.74]. There are two major sets of rules, Japanese and Chinese; other variants also exist. The Japanese method of scoring is by territory, the

Chinese method of scoring is by area. The American Go Associations recently adopted a set of rules [AGA] that allow tournament players to select either method of scoring before each game. As elaborated by Ikeda Toshio in his "The Rules of Go" [IT], the Japanese rules use a number of complex, occasionally ambiguous precedents to decide when a game ends and how to score it. Ikeda proposes a simplified rule set which has the flavor of the Chinese set of rules. The question of when and how a game ends is precisely answered in the rules of Mathematical Go.

The rules of go, Japanese version: overview

This description of Japanese rules was adapted from several articles appearing in Go World [GW]. Go World is the major English Language publication on go, and periodically summarizes the prevailing Japanese official rules of the game.

(A) Go is a game played between two players, Black and White, upon a go board using black and white stones.

(B) Black moves first.

(C) Players alternate moves.

(D) Each move consists of either placing a stone of one's own color on a vacant intersection. or passing. By agreement of both players, Black's first move may consist of placing a specified number of "handicap stones" at designated locations.

(E) Passing is always legal.

A move is legal if both:

(F1) It does not repeat the board position two moves earlier, and

(F2) It does not capture itself (commit suicide).

(G) Opponent's stones are captured when their last liberty is taken. (Liberties and grouping of stones are defined below.) Captured stones are removed from the board and placed in a separate pile used for scoring after the game is over. Removing captured stones is the only time stones placed on the board are moved. Captured stones are also referred to as prisoners.

(H) The game ends after two consecutive passes.

(I) The game ends in a draw when in a repetitive position neither player "gives way." See below for further discussion.

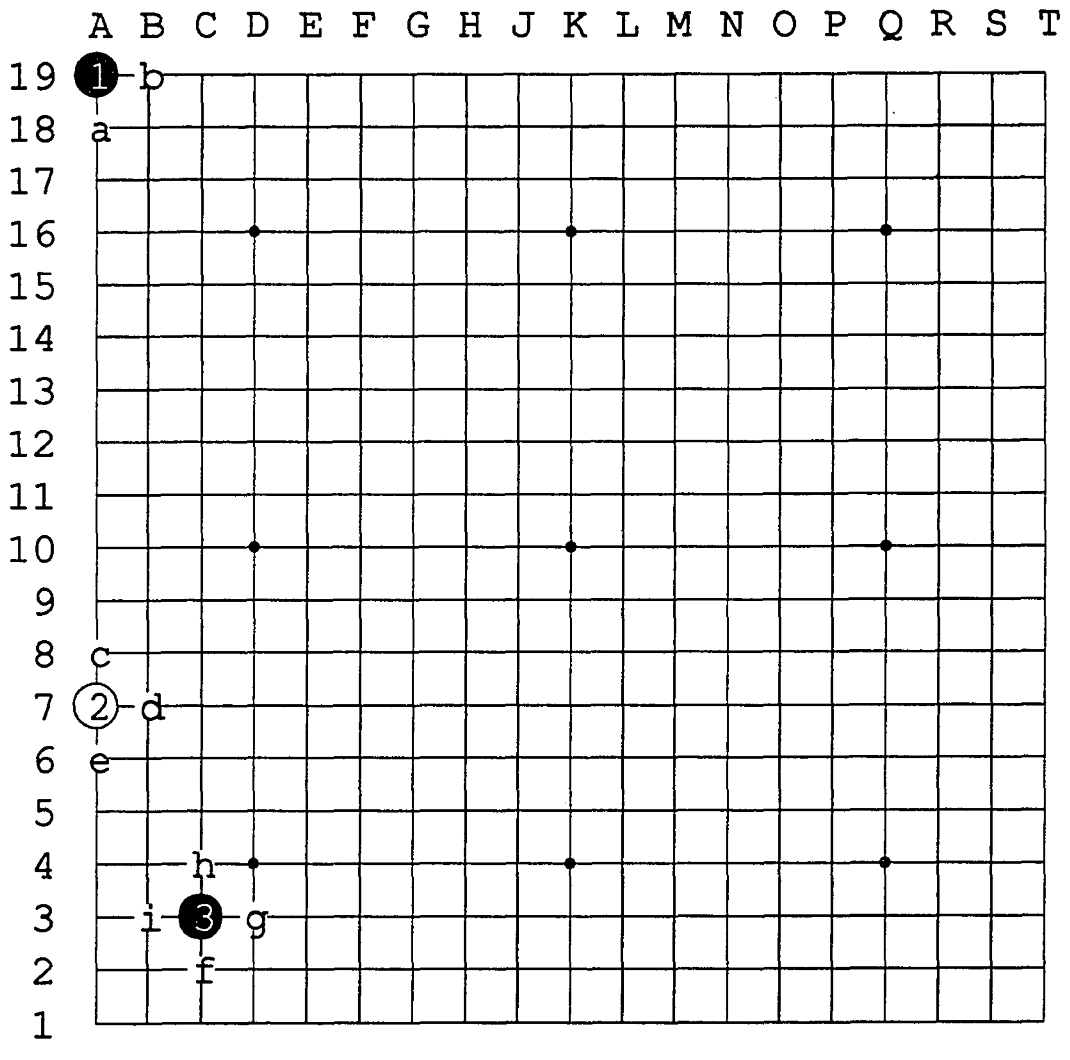
(J) If after the game is over, there is a dispute over the scoring, a referee can adjudicate the position, or there can be a "resolution session."

(K) Each player's score is determined by the number of points he has surrounded minus the number of prisoners his opponent has captured. The player with the largest score wins. The score is adjusted by any previously agreed upon point handicap, or "komi." In an even game, White is usually given 5.5 points to compensate for Black having the first move. The result of a game is either Win/Loss or Draw/Draw.

The Japanese rules of go: details

The Equipment

Go is played on a board marked by two sets of N parallel lines orthogonal to each other. See Figure 3.1. Each player, Black or White, has a set of stones of his own color, of unlimited supply, and a location for storing captured opponent's stones.


A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
1
2
3
a
b
c
d
e
f
g
h
i

A corner, side, and middle play.
Figure 3.1

The graph formed by the lines on the board is an N x N array of intersections (henceforth called points). A straight line between two neighboring points defines those two points as connected. The geometry of a rectangular board partitions the points into 3 types: 4 corner points, 4 x (N - 2) edge points, and (N-2) x (N-2) central points.

For descriptive purposes, the points are labeled as coordinates of a graph, column followed by row. The columns are named from left to right using A - H, then J - T. (I is omitted because of the potential confusion with 1 or J.) The rows are labeled from bottom to top with the numbers 1-19. Alternatively, the stones are numbered as played (move number), with empty points under discussion labeled directly by lower case letters.

The 19 x 19 board is used for regular competition. The 9 x 9 and 13 x 13 boards are frequently used by beginners and novices. Professional players have experimented with serious matches on the 21 x 21 board. Go World [GW] has published a series of end game puzzles highlighting the peculiar characteristics of the 6 x 6 and 7 x 7 board called "Go in Lilliput." Several papers treat the complexity of the game as a function of the board size. The rules are the same for any size board. The complexity of the game tree and the richness of the strategy vary dramatically with size.

Black's First Move

The 19 x 19 go board has 9 intersections marked with enlarged dots: D4, D10, D16, K4, K10, K16, Q4, Q10, and Q16. When two players of unequal strength wish to play an even contest, the weaker player as Black makes his first move by placing several handicap stones on the board according to the following table.

2 stone handicap: D4, Q16.

3 stone handicap: D4, Q4, Q16.

4 stone handicap: D4, D16, Q4, Q16.

5 stone handicap: D4, D16, K10, Q4, Q16.

6 stone handicap: D4, D10, D16, Q4, Q10, Q16.

7 stone handicap: D4, D10, D16, K10, Q4, Q10, Q16.

8 stone handicap: D4, D10, D16, K4, K16, Q4, Q10, Q16.

9 stone handicap: D4, D10, D16, K4, K10, K16, Q4, Q10, Q16.

After Black's first move the game proceeds in normal fashion.

Liberties, Captures, Chains, Groups, Armies, and Legal Moves

Eskimos have scores of words to describe snow; those living in temperate climates have but a few. The language for describing when collections of stones and when they may be captured, either immediately, eventually during the main game period, later during the resolution period, with or without a certain expenditure of moves, might be plentiful in Japanese, but could use a few more terms in English. The existing terms are often imprecise.

In his early book "Go and Go-Moku," [LE] Edward Lasker used the terms chains, groups, and armies to loosely describe aggregates of stones of the same color. We shall use Lasker's terms in a more rigorous fashion.

A chain is the unit of capture. A group is the unit of life. An army is a loosely knit connection of chains and groups with the potential to form one large group. A liberty is an empty point connected to a chain.

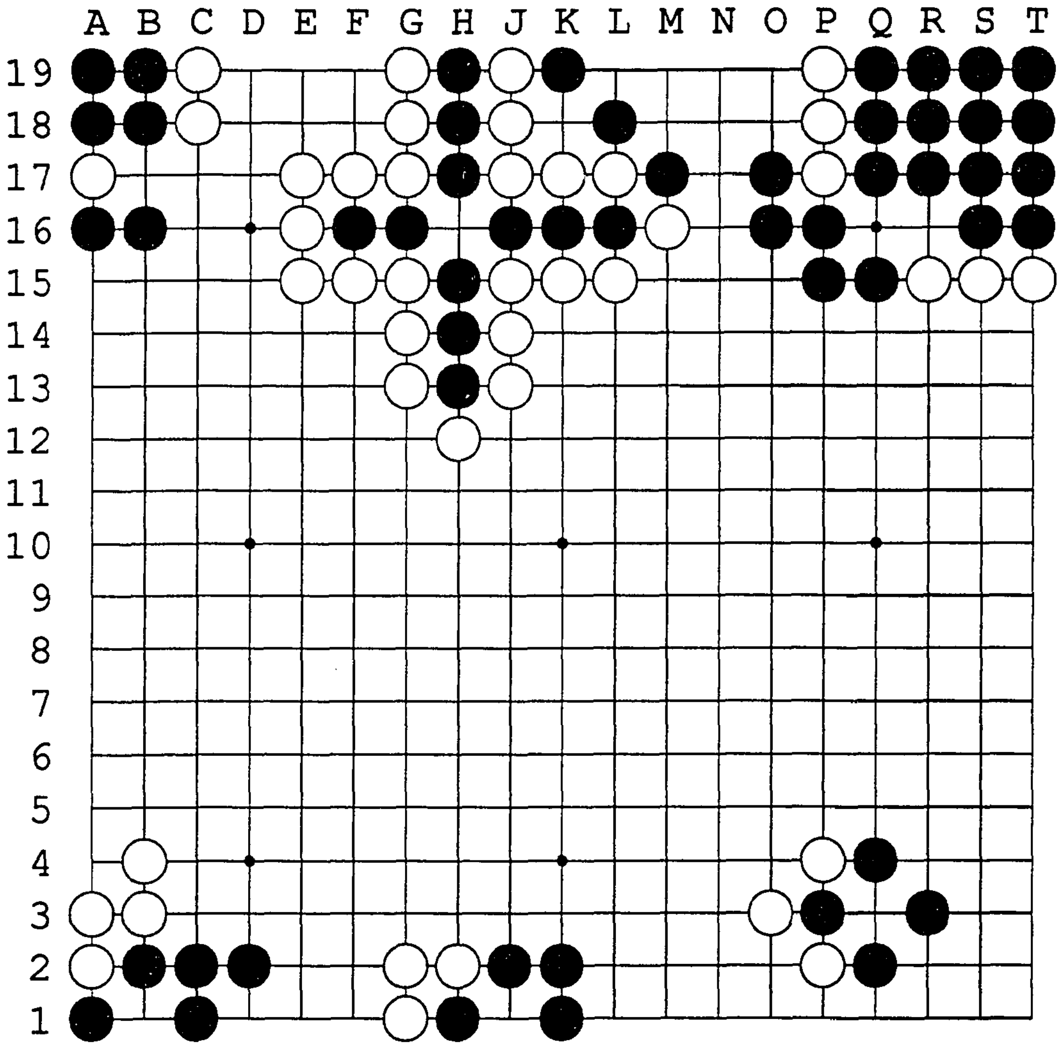
When stones of the same color are connected by horizontal and/or vertical lines then these stones form one chain. The liberties of a chain are the distinct liberties of each member of the chain. When a move reduces the number of liberties of a chain (or chains) of an opponent to zero, then the chain(s) is (are) captured. See rule (G).

Figure 3.1 shows the 19 x 19 board, columns and rows labeled, and the results of a game after 3 moves (definitely not good play). Black 1, played at A19, is on a corner point with two lines connecting it to its liberties A18 and B19, labeled a and b. White 2, played at A7, is on an edge point with three lines connecting it to its liberties A8, B7, and A6, labeled c, d, and e. Black 3 played at C3 is on a central point with four lines connecting it to its liberties C2, D3, C4, and B3, labeled f, g, h, and i.

Rule (F2) means that a legal move must have at least one liberty. A capture removes some opponents stones, so capturing moves are legal unless they violate rule (F1).

Thus, in Figure 3.1, White plays at a and b would capture the Black stone labeled 1 at A19; Black plays at c, d and e would capture the White stone labeled 2 at A7; and White plays at f, g, h, i, and j would capture the Black stone labeled 3 at A3.

Figure 3.2 shows more examples of capturing moves. Black at B17 captures one White stone and connects two chains into one chain. White at B17 captures four Black stones and


A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

Capturing positions. Ko positions.
Figure 3.2

nearly connects two chains. In the upper middle position, Black on play cannot play at H16 because it would connect four chains, each with one (shared) liberty into one chain with no liberties, an illegal move, by rule (F2). Black can play at K18, capturing five White stones and creating several liberties for two of the threatened chains. In the original position, White at H16 captures four chains, eleven stones. If Black had played first at K18, White at H16 captures two chains of five stones. Notice that in the original position, White can play at H16 which has no immediate liberties until after the capturing stones are removed.

On the upper right is an example of an immediate recapture that is legal. Black has one chain of fourteen stones with two liberties, and another chain of five stones with seven liberties. By playing at Q16, he connects the two chains to form one chain of twenty stones with seven liberties. White can play at Q16 threatening R16 and capture of fourteen stones. If Black plays at R16, he captures one White stone but reduces his chain to only one liberty. White can replay at the same Q16 location and capture fifteen Black Stones because the resulting board position does not repeat itself.

Rule (G) says that captured stones are immediately removed from the board. Roughly speaking, a "live group" is a chain or aggregate of chains that cannot be captured with "normal play." A "dead group" is a chain or aggregate of chains that cannot be defended from capture with "normal play." An "unsettled group" is a chain or aggregate of chains whose life and death status is not yet decided. The stones of a strategically "dead group" are not removed from the board until they are actually captured.

Armies are a collection of chains or groups that are near each other with the potential to form one living group. Roughly speaking, the opening phase of go stakes out armies loosely across the board. The middle phase of go proceeds to settle these armies, to resolve their life or death status, and to extend from living armies to enlarge territories. As the game proceeds, unsettled groups become settled, and the game ultimately ends. After the game ends, groups (chains) that are dead are removed from the board and become prisoners of the opponent just as if captured during play.

Ko

Rule (F1) says that it is illegal to play a move that would repeat the board position of two moves earlier. This position is known as "ko." When a player captures a single stone with a move such that the capturing stone has exactly one liberty, then and only then would a recapture reproduce the position of two moves earlier. That capture is illegal and cannot be played. Figure 3.2 shows examples of a ko position in the corner, on the side, and in the center. White to play can capture at B1, J1, or Q3. If White makes any of these moves, Black must wait at least one move before recapturing.

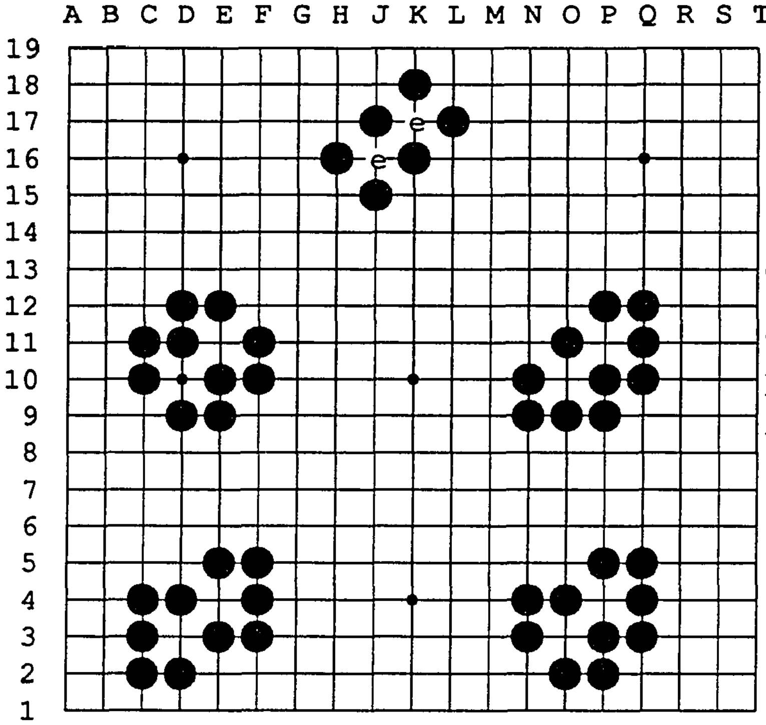
The settling of a large group may depend on ko. The move a player makes immediately after his opponent makes a ko capture is known as a "ko threat." If player 2 ignores a ko threat and fills (connects) the ko, then player 1 gets to make two moves outside the ko area without player 2's intervening move. If player 2 responds to the ko threat and fails to fill the ko, then a symmetrical strategic position has arrived with colors reversed. The sequence: <P1 takes ko, P2 makes ko threat, P1 replies, P2 takes ko, P1 makes ko threat, P2 replies> repeated until <Px takes ko, Py makes ko threat, Px fills ko> frequently occurs in actual play and is known as a "ko fight." There is a whole taxonomy of ko positions depending on how many moves are needed by each side to ultimately resolve the position.

Two eyes make life

Liberties can be thought of as internal and external. Internal liberties are surrounded by stones of their own color, and optionally by sides of the board. Internal liberties form "eyes." External liberties are liberties which are not internal liberties.

A group with two distinct eyes of size 1 is forever immune from capture because after the opponent fills all the external liberties, he cannot legally fill either eye. However if part of the external wall is insecure and can be captured, then the group may die. An eye (internal liberty) not securely surrounded is called a "false eye."

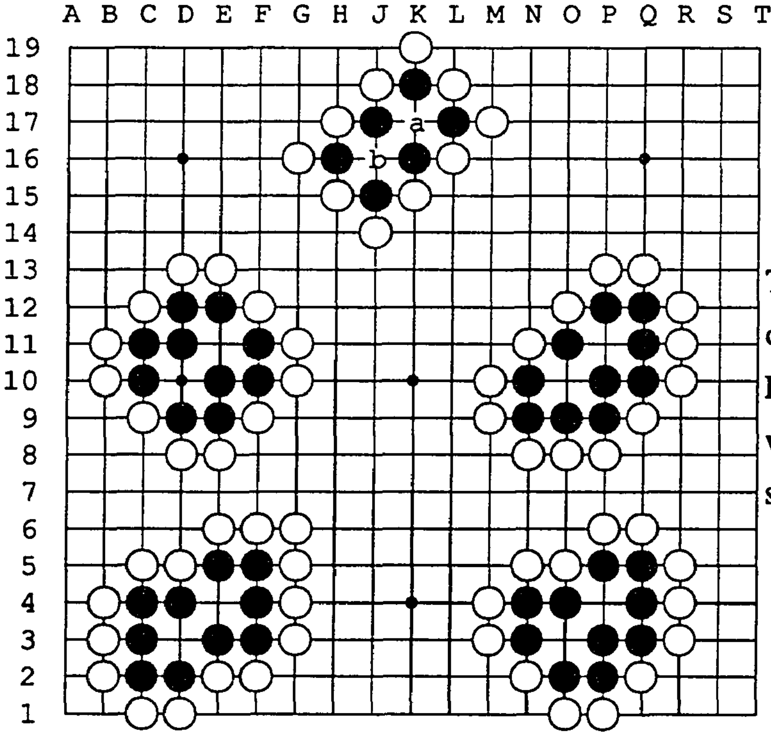
A live group may contain many chains, but each chain must share at least two of the group's eyes. (Living groups of minimal size are often used in Mathematical Go problems, see Chapter 7.) The minimal number of stones needed to surround two central eyes is ten. There are four distinct patterns that can do this. (From here on, distinct means ignoring rotation, reflection and translation.) Figure 3.3 illustrates these four patterns. The most economical surrounding of two eyes is shown in the upper center. "e" marks two empty spaces (internal liberties, eyes). The six Black stones have ten external liberties, and only two of the six chains (stones) share both eyes. By adding four stones in the indicated four patterns, one joins six chains into two chains, each sharing two eyes. These are shown in the middle and bottom of Figure 3.3.


A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
e
e

Two eyes in the middle of the board

Four live Black groups using minimal stones (10).

Figure 3.3


A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
a
b

Two eyes in the middle of the board.

Four live Black groups. with encircling White stones.

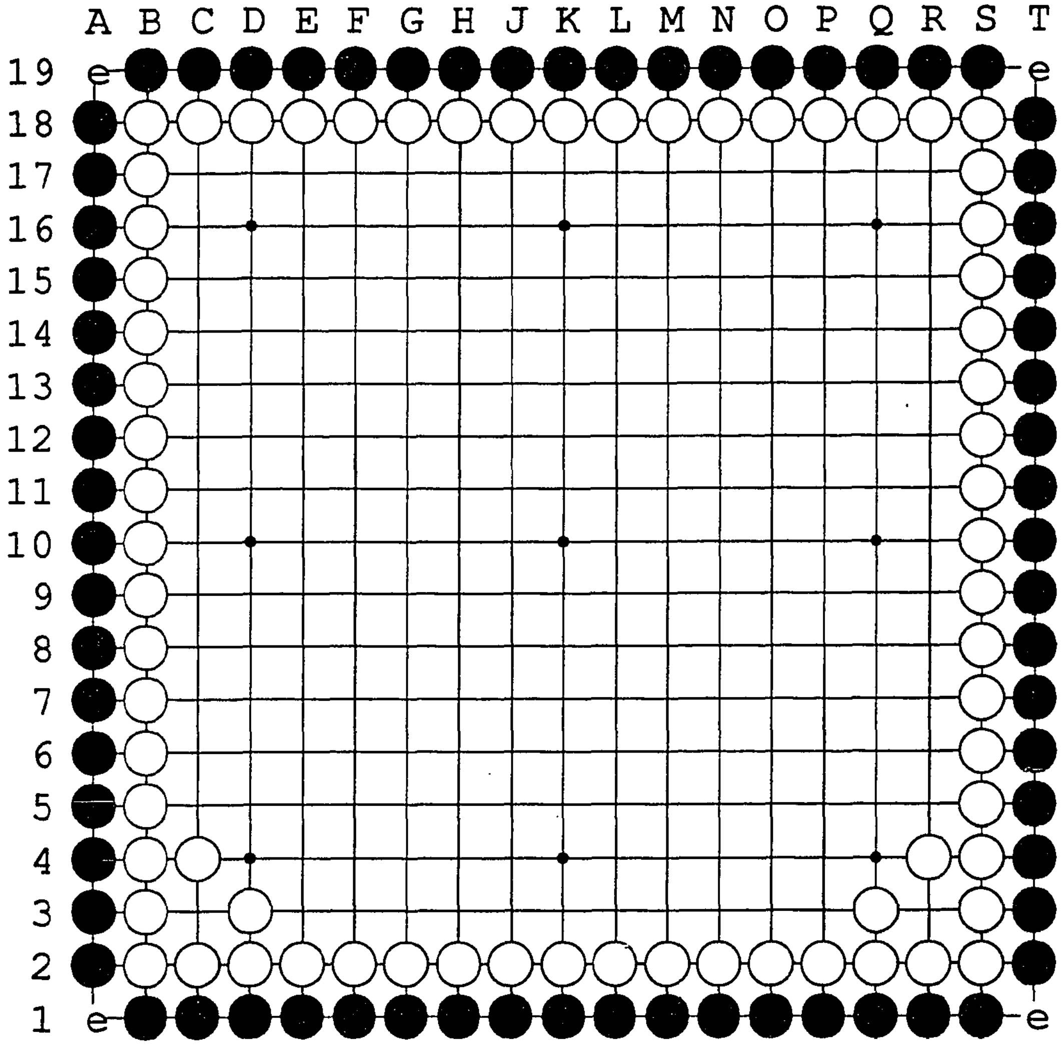
Figure 3.4

The operational test of a live group is this: can an opponent fill all the external liberties and then capture part (or all) of the group by filling in eyes? Figure 3.4 demonstrates that all of the lower four patterns are immune from capture, and they are alive; however, the top pattern is not. After White completely surrounds each lower group, White can't fill any eye, since that play would have no liberties and not capture any chain(s). Notice that in the upper central position, by playing at a or b White captures two stones, and can capture four more stones on the next move. It is interesting that every one of these patterns uses at least two chains to form a group, rather than just one. Figure 3.5 shows an even more curious living group [LA]. Black has four chains, each of 17 stones, each chain sharing two of the groups four eyes. If any Black stone were replaced by a White stone, then Black has a dead group as it stands. ("As it stands" means no additional play until the game is over is needed to establish the group's status.)

To continue enumerating live groups with minimal stones, there are three distinct corner patterns, and four distinct edge patterns. These are illustrated in Figure 3.6. Notice that five of these seven patterns have two chains.

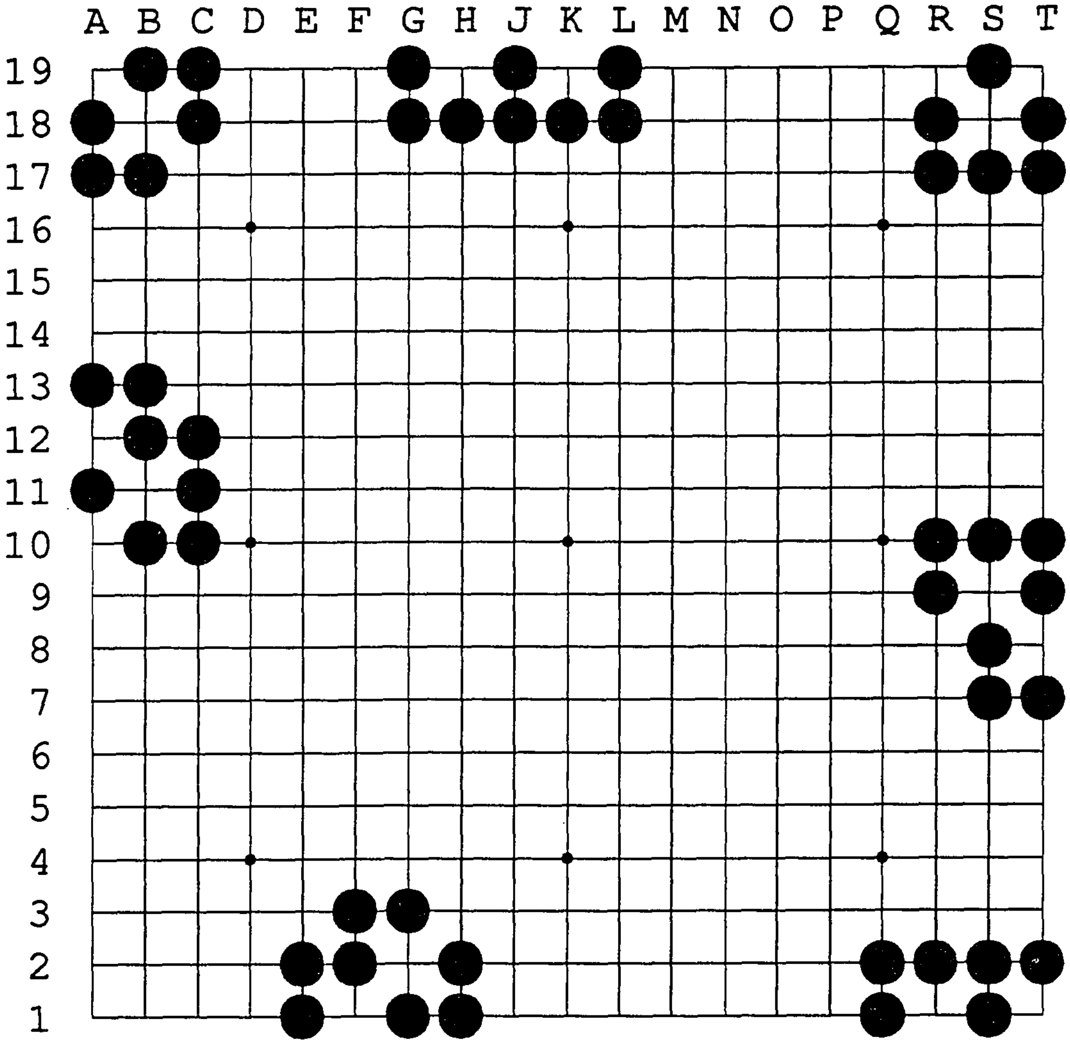
Big eyes, or "nakade"

An eye is one or more internal liberties surrounded by stones of one color. An eye of size greater than two may or may not be able to form two eyes, should the need arise. Figure 3.7 illustrates these points. In the upper left, the Black group is dead as it stands; it can never partition two liberties in a row into two eyes. If this were a resolution session and White had to kill Black, he could play at A16, threatening immediate capture, Black might capture the in-

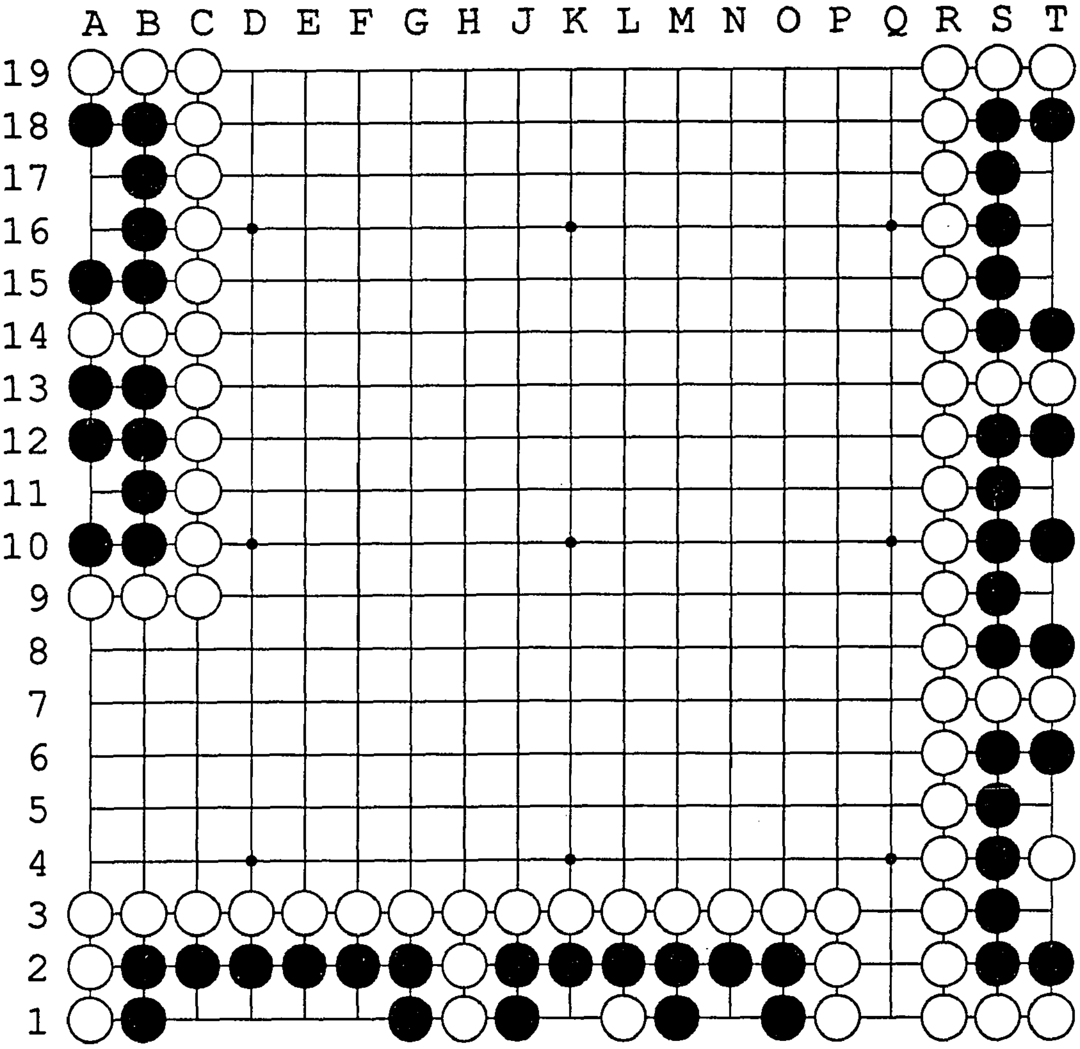

A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
e
e
e
e

A black living group with four eyes, and four chains. Replace any Black stone by a White stone and Black's group is dead as it stands.

Figure 3.5


A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Three live corner groups, minimal stones (6).
Four live groups on the side, minimal stones (8).
Figure 3.6


A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

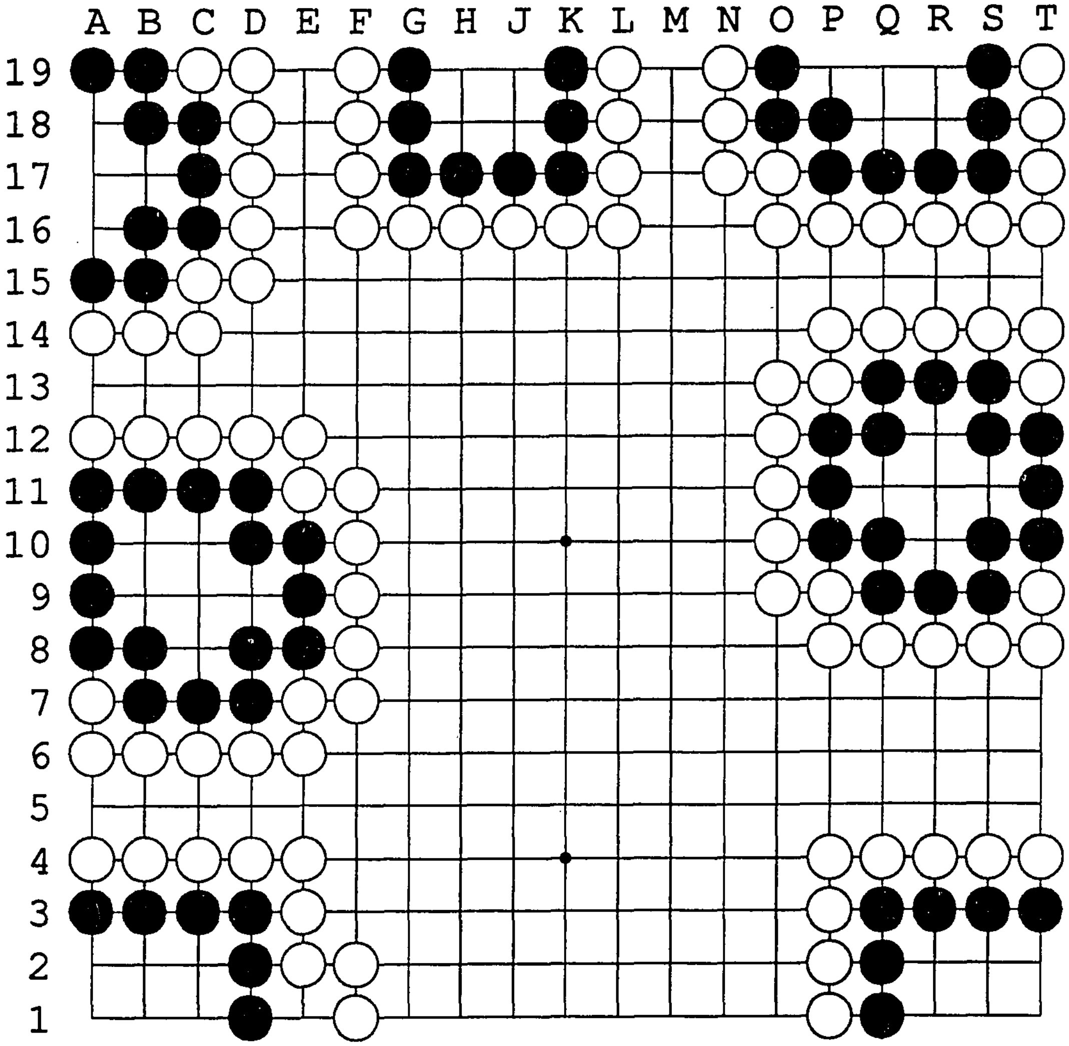
Two, three, and four liberties in a row.
Figure 3.7

vader with A17. The position would now look like the group from A9 to A13 on the left where White can capture immediately.

The Black group on the upper right has three liberties in a row. It is unsettled. Black to play occupies the middle point, T16, and it is clearly alive. (See the middle right pattern.) White to play also occupies the middle point and the group is dead as it stands. (See the lower right pattern.) If this were a resolution session and White had to kill Black, he could play at T3 threatening immediate capture. Black might respond at T5 capturing the two invaders, and the pattern is the same as the upper left pattern.

The lower left Black group is alive as it stands. Should White try to kill Black by playing at D1 or E1, Black can play at the other of the two points. (See the bottom right pattern.) It is now illegal for White to play at either of the two remaining eyes, so Black lives, and the White invader is a dead stone. This pattern is different from the middle right pattern because it provides a ko threat for White. If a ko fight were occurring elsewhere on the board, and Black had made a ko capture, hoping to fill the ko square on the next move, White should play at D1 (or E1). If Black fills the ko square, then White can now play at E1 and the Black group is dead as it stands. In a resolution session, after White has played at both D1 and E1 White kills Black by playing at C1, Black captures three stones with F1, and White plays at D1, in the center of three liberties, reproducing the lower right pattern.

There exists a finite number of one eye shapes that are not alive as they stand. They are known as nakade. Figure 3.8 illustrates the rest of these shapes. The upper center pattern is dead as it stands. The upper and middle left and right patterns are each unsettled. The vital points for each side to play are A17, C9, Q19, and R11. These five nakade are the same


A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

Five nakade and two corner positions.
Figure 3.8

whether in the corner, the edge or the middle.

There are a number of shapes in the corner, not called nakade, whose status is unsettled. For example, the two by three rectangle is alive in the center or edge, regardless of external liberties. In the lower right, with no external liberties, the shape is unsettled. Black plays first at S1 or S2 and lives as it stands. White plays first at S2, Black's best response is S1, White plays at T2, and Black's response of R2 fails, Black is dead. In the lower left, Black plays first at B2 or B1 and lives. If White plays first at B1, then Black's best response is at B2, White plays at C1. If Black plays at A1, White plays at A2, capturing A1 in ko. Were two external liberties available, Black would play C2 followed by A1 and live unconditionally. Since C2 allows White to capture at E1, Black must win a ko fight over the A1 and A2 squares before playing at C2. Black can fight the ko now or wait for White to approach at E1.

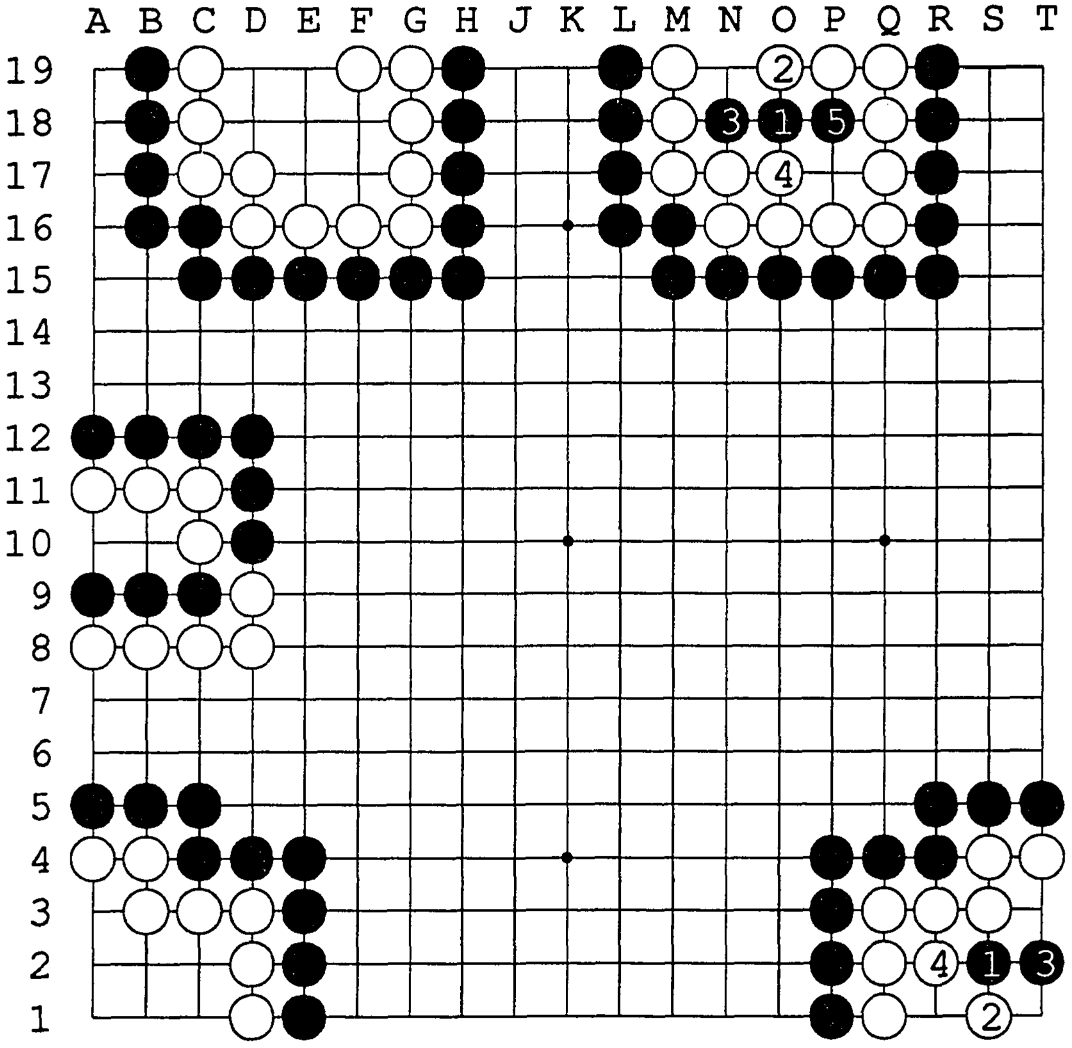
Neutral points, or "dame" (pronounced dah'may)

When the game ends, most of the board is settled into areas of territory for one or the other player. Vacant points left in between these areas are neutral points, or dame. They have no effect on the score. In friendly games, among players anxious to play another game (or to tally the score of the current game), it is perfectly acceptable for one side to fill in all the dame with his color.

Seki

Groups with two eyes are alive. Groups with no eyes or one eye often live by capturing prisoners. When each side tries to capture an opponent's chain(s) in order for its own chain(s) to live, it is called a capturing race. When a capturing race is such that neither side can capture the other without being captured first, the condition is called seki. Both sides' stones remain on the board after the game ends, and neither internal nor shared liberties are counted as territory for either side. Figure 3.9 illustrates these cases. In the middle left, the White chain of A11, B11, C11, and C10 has two liberties, A10 and B10. The Black chain of A9, B9, and C9 shares the same two liberties. If either side plays at one of the shared liberties, it reduces the opponent's chain to one liberty, but also reduces itself to one liberty. On the next play the opponent could capture the attacker.

Of course it is legal to sacrifice one's stones, but never forced. Should Black play A10 and White reply A11, Black loses eight points (four prisoners and four points of territory), but Black can use this as a ko threat if the ko fight were worth more than eight points. On the upper left is a seven point shape subject to seki. The position is repeated on the upper right with moves one through five shown; best play for both sides, the result is a seki. White's nine and Black's three stones stay alive on the board at the end of the game; and N19 and O17 belong to no one. The lower left pattern is similar. The position is repeated on the lower right with moves one through four shown, best play on both sides. The result is seki. A White play at T3 results in immediate capture; a White play at R1 leaves White dead with a nakade previously discussed. A White play at T1 is followed by Black at R1, White at S1, and White has lost a point and made no progress.


A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
1
2
3
4
5
1
2
3
4

Seki positions.
Figure 3.9

No result when neither side gives way

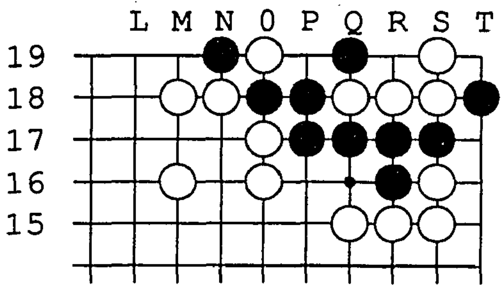
It is illegal to repeat a position of cycle length two, but it is legal to repeat a position of cycle length greater than two. An example of a cycle of length four moves, called "endless life," is illustrated in Figure 3.10. If White plays elsewhere at move 2 or move 4 (a move played elsewhere is called "tenuki"), Black will form a live group for his stones with four or five points of territory and four prisoners, for a local gain of about nine points. If Black plays elsewhere at move 1 or move 3, White will capture ten Black prisoners with roughly 18 points of territory; for a local gain of 28 points. If either side felt it would still be ahead over the entire board after playing tenuki, it might break the cycle. Otherwise, the game is declared no result.

An example of a cycle of length six moves, called "triple ko," is illustrated in Figure 3.11. By playing tenuki, a side will cede to the opponent 11 prisoners and 14 points of territory. As mentioned above, unless either side felt it would still be ahead over the entire board after playing tenuki, the game is declared no result.

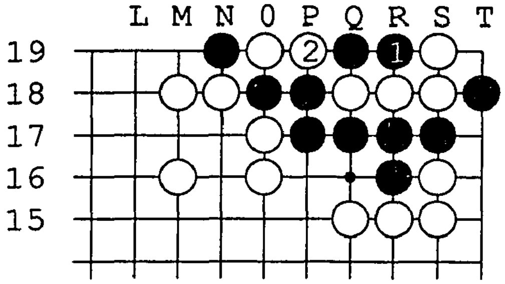
Resolution session

As mentioned above, Ikeda illustrates how the Japanese rules are complex because they adjudicate positions by precedent, requiring a knowledgeable referee to be present. James Davies, in "The Japanese Rules of Go:" [DJ] discusses how a game ends when no referee is present:

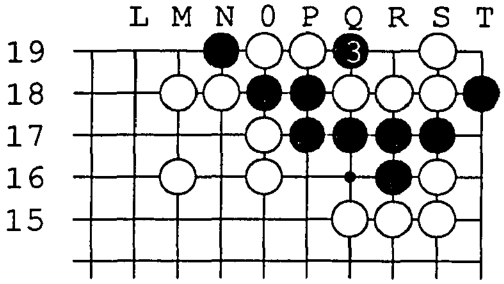
> What most rule theoreticians seem to do is allow passing and let two consecutive


L M N 0 P Q R S T
19
18
17
16
15

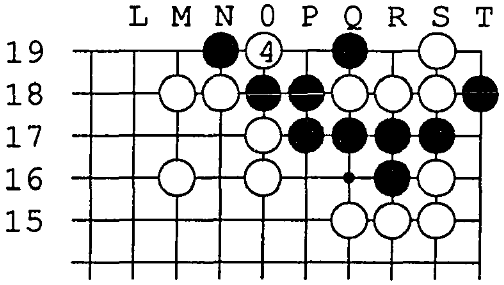
Black to play and not die


L M N 0 P Q R S T
19
18
17
16
15
2
1

Black 1 sacrifices one stone. If White replies elsewhere Black lives unconditionally by capturing the group of four White stones with a move at T19. White 2 captures two Black stones, threatening to capture N19 with M19 and kill the whole Black group.

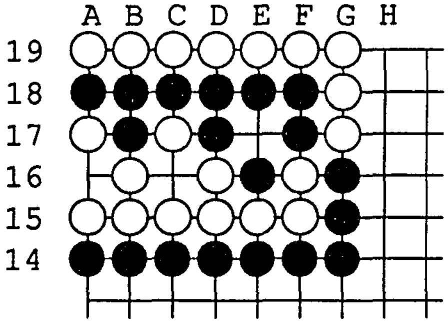

L M N 0 P Q R S T
19
18
17
16
15
3

Black 3 captures two White stones. If White plays elsewhere Black lives unconditionally by playing at O19 for one eye. His second eye is secure since he wins the capturing race and kills White's group of stones on Q18, R18, S18, and S19.

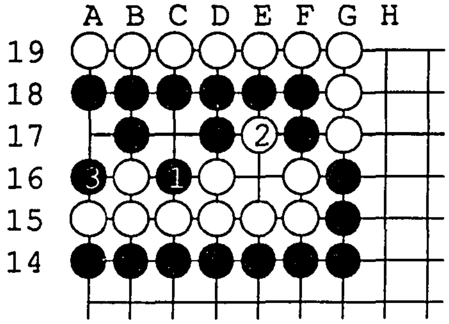

L M N 0 P Q R S T
19
18
17
16
15
4

White 4 stops Black's second eye and reproduces the initial position. If neither side gives way, the game is "no result."
Note: if Black 1 at P19, White 2 at R19 reduces Black to one nakade, Black is dead. See Figure 3.8, upper right, with its discussion.

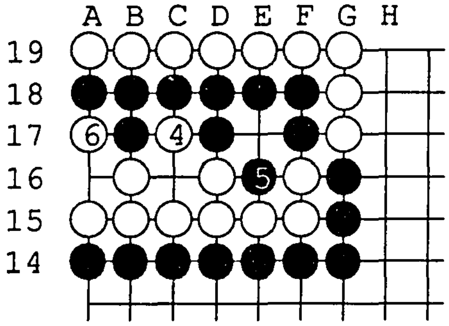
An example of "Endless Life" [KY, p.30]
Figure 3.10



Black to play and not die
Black's chain of eight stones, mostly on row 18, has only one liberty, i.e. is in immediate danger of capture.



Black must capture one of the two kos. Assume he plays at 1, capturing C17. White now has one liberty. Each succeeding play threatens to kill the opponent. Each play is now forced. White plays at 2, capturing E16. Black plays at 3, capturing A17.



White plays at 4, capturing C16, Black plays at 5 capturing E17. White plays at 6, capturing A16. The original position is restored with Black on play. If neither side gives way, the game is "no result."

Triple ko, a cycle of length six.
Figure 3.11

> passes define the end of the game. They also allow (require, in one case) play to continue after the end of the game to settle life-and-death questions. To keep things fair during this second phase of the game, if a player passes he must pay his opponent one penalty stone, and the number of stones played or paid by both sides must be the same.

The resolution phase ends with another set of two consecutive passes. All stones remaining on the board are alive.

Before the end of the game, passing is not accompanied by paying a penalty stone. The method of scoring is such that playing a move within your own territory without affecting any group's status, (your own or your opponent's), loses one point. Placing a stone within your opponent's territory without affecting any group's status, likewise loses one point. By passing, a player neither loses nor gains any points. Thus it is strategically wrong (not illegal) to try to kill an opponent's already dead group within one's own territory, or to try to rescue your own dead group within an opponent's territory. After the end of the game (in a resolution session), since it is a point penalty to pass, the disincentive to settle each life and death question out to a conclusion is lifted.

Chinese rules

Currently there are a number of Chinese versions of the rules of go espoused by various organizations. What follows has been culled from tournament announcements placed on the INTERNET bulletin board, rec.games.go [RGG]. The main differences between Chinese and Japanese rules are:

a) The komi often varies among different organizations, $7\frac{1}{2}$ or even $9\frac{1}{2}$ have been used, although $5\frac{1}{2}$ is the most common.

b) Each player's score is determined by the number of his stones on the board and the number of points surrounded (territory). This implies:

i) Dame (Japanese neutral points) count. Ko fights for the last dame are meaningful.

ii) Playing in one's own territory carries no penalty.

iii) One should play the game out until the only thing left is settled territory. There is no disincentive to pass early, as under Japanese rules. There is no long list of special situations as under Japanese rules.

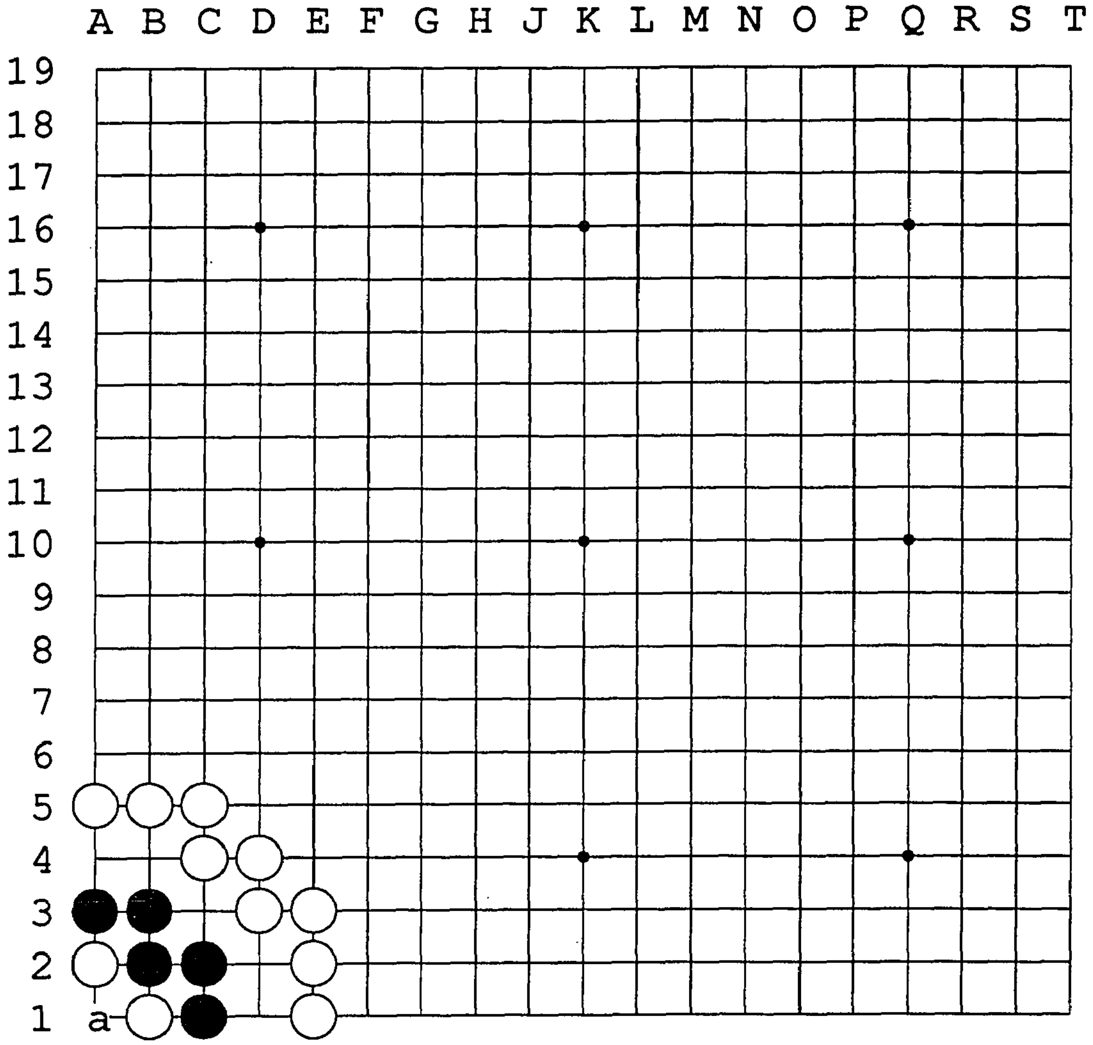
c) Suicide is permitted when the suicide stones number greater than one. See Figure 3.12 for an example where suicide can generate an effective ko threat not available under Japanese rules.

d) Liberties in a seki position are counted. Assessment depends on the ratio of stones on the perimeter of the area in question. There are several variations to this rule, and some organizations do not apply it.

e) Black may place his handicap stones anywhere on the board, ignoring the fixed points designated under the Japanese rules.

Rankings

There are local, national, and international rating scales. The scale from lowest to highest is:


A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
a

The availability of a play at "a" is a potential ko threat that White may use at any time. When White moves at "a", he removes three stones at A1, A2, and B1 as Black's prisoners. Now, a followup play by White at "a" will kill Black.

Figure 3.12

50 kyu decrementing by steps of 1 to 1 kyu.

1 dan amateur is stronger than 1 kyu.

1 dan amateur incrementing by steps to 6 dan amateur.

1 dan professional is stronger than 6 dan amateur,

1 dan professional incrementing by steps to 9 dan professional.

A novice who has just learned the rules has strength between 35-50 kyu, depending on his familiarity with games in general. After playing seriously for a few weeks, one's strength should be at 15-20 kyu. The ranking of the best computer go programs is hard to establish, but has been estimated at 10 to 15 kyu.

Between non-professionals, each level of difference can be compensated for by one handicap stone. Between professionals, each three levels of difference can be compensated for by one handicap stone.

Mathematical Go

Mathematical Go may be thought of as a class of composed endgame problems in the traditional game of go. Berlekamp [BE2, p.77] calls the rules for Mathematical Go "Basic Mathematical Go (ancient Chinese Dialect)," but they correspond more closely to modern Japanese rules than to modern Chinese rules. Using the same numbering scheme as before ((A) through (K)) we describe the similarities and differences. We add rules as needed.

(A) Mathematical Go is a game played between two players, Black and White, upon a go board using Black and White stones, as in go.

(B) The initial position of a particular game is set up by the problem composer, the player to move is specified, as in go endgame problems.

(C) Players alternate moves, as in go.

(D) Each move consists of either placing a stone of one's own color on a vacant intersection or giving one's opponent a prisoner. One must have captured the prisoner, unlike the Japanese resolution session.

(E) Passing without transferring a captive to the opponent is illegal, unlike Japanese go before the game ends.

(F1) There is no ko restriction, unlike any other version of go. This means immediate recapture in a ko position is legal.

(F2) Suicide moves are illegal as in Japanese rules of go.

(F3) Immortality By Convention [BE2, p.76]:

> Mathematical Go is very concerned with the study of partial game positions, or more simply, "games." Typically, each of these games contains a region of only a few vacant nodes, which are surrounded by occupied nodes. When investigating the play within such a region, we assume that the stones on its boundary are "immortal."

Once a group is declared immortal, one can (and will eventually) fill its eyes. This is unlike go, but produces similar outcomes. Effectively, it counts territory inside living groups by decrementing one with each move.

(G) Captures are defined as in go.

(H) The game ends when a player has no legal move, and the player to move loses. This definition is unlike go, but produces similar outcomes. It is completely unambiguous.

(I) The game ends in a draw when in a repetitive position (including a ko position) neither player "gives way." For cycles greater than length two, this is as in go.

(J) Because of the way the end of a game is defined, it is impossible to have a dispute in the scoring. There is never a need for adjudication of a position nor for a resolution session.

(K) There is no score of a game. As stated in rule (H), the player unable to move loses. In Chapter 6 we discuss the relationship between the outcome of a game under Mathematical Go and go scoring.

(L) Games that end must end in a win or loss. In go, there is a draw by no result when a repetitive position occurs, and there is also a draw by score ("jigo" in Japanese) when the method of scoring produces a tie (either equal area or equal territory plus captives).

This chapter has described how to play go and Mathematical Go. The next chapter will use these rules on a real example.

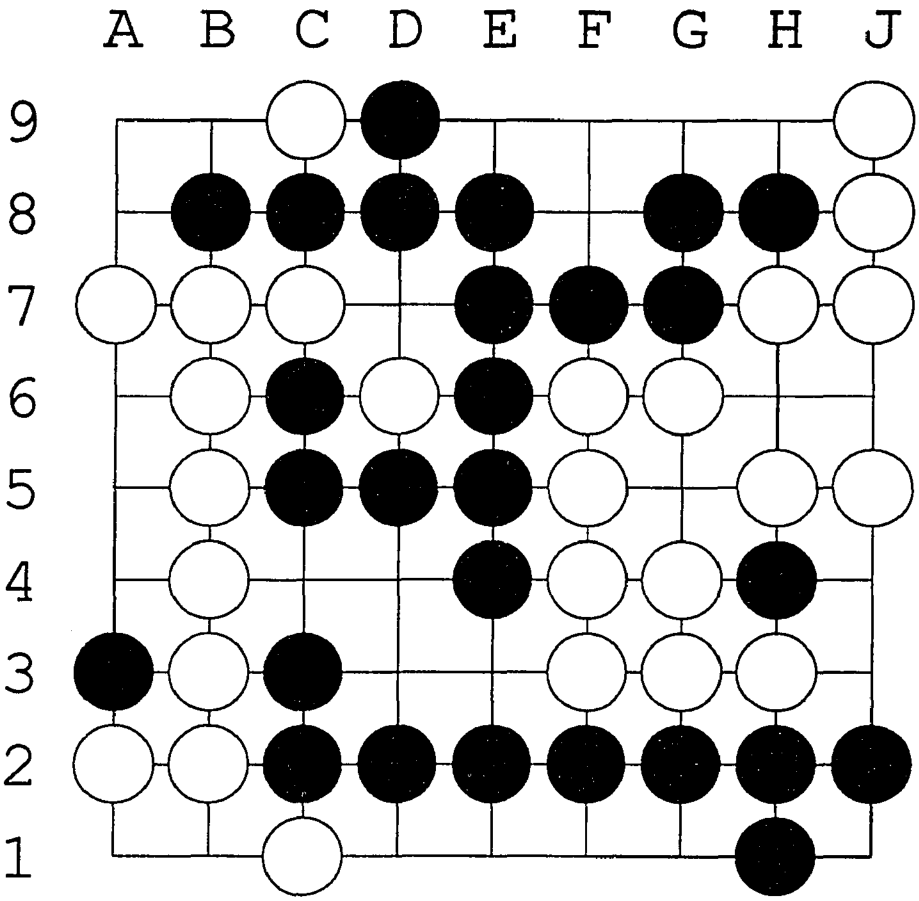
CHAPTER 4. A Mathematical Go problem and its solution

Berlekamp's paper "Introductory Overview of Mathematical Go Endgames" [BE2] presents one Mathematical Go problem and its solution. It is a very good problem because it touches on most of the concepts in Mathematical Go. It was the only example of a Mathematical Go problem published prior to the publication of David Wolfe's thesis "Mathematics of Go: Chilling Corridors" [WD1] in late 1991. (Berlekamp was Wolfe's thesis advisor.) This chapter discusses in detail Berlekamp's problem.

Figure 4.1 is a representation of Berlekamp's problem. The solution of this problem shows how Mathematical Go has extended the theory of WW to positions with potential repetitions. It also introduces a well defined theory of winning certain very close go positions.

In this problem, the Black stone at A3 is a prisoner, the Black stone at H4 is a potential prisoner, and all the other Black stones are an army that can't be prevented from forming one living group. White has two potential prisoners at C9 and C1, and two live groups, one on the right, and one on the left. White forms two eyes on the left by playing at A4 or A5. White forms two eyes on the right by playing at H6 or J6.

Certain features of this example are true for all Mathematical Go problems. Mathematical Go problems are always whole board problems. The strategy treats different areas of the board as totally separate from each other. The only moves are what in traditional go would be called "reducing" moves. The areas subject to reduction are always part of live armies. There is never an unsettled group present which has to fight for its life or "seki." There is never a killable group to attack. There is never part of one group sacrificed for another. There is never a


A B C D E F G H J
9
8
7
6
5
4
3
2
1

White to move and win

A whole board problem adapted from Berlekamp, "Introductory Overview of Mathematical Go Endgames" [BE2, p.75]

Figure 4.1

sequence of plays that hinge on the tactic of connecting a floating group, or merging two groups with with one eye into one live group, or leaning on one group to gain strength to attack another. For a marked contrast with Mathematical Go, traditional go strategy says stones may have power and influence that act at a distance, radiating throughout the board. The Appendix to this thesis offers two examples.

The solution to Berlekamp's problem and to all Mathematical Go problems involves four major steps:

I) Partition the position into its significant tiles.

II) Simplify the value of each tile as much as possible.

III) Perform a mathematical mapping on each simplified tile from the basic WW value to a new value. This mapping is called chilling, and will be discussed in detail in this chapter. After chilling, more simplification may occur.

IV) Using the chilled values, play out the game using WW techniques. This involves creating a partial order of the best move from each tile and selecting the move with the greatest incentive; this largest increment of value (or least decrement) will be the move most beneficial to the player.

The solution to Berlekamp's problem shows that White will win by one point with best play on both sides. It shows that if Black offers appropriate resistance (as pointed out in Berlekamp's paper) then White must find unique winning moves at moves 1, 3, 5, 9, 11, 13, 15, and 17.

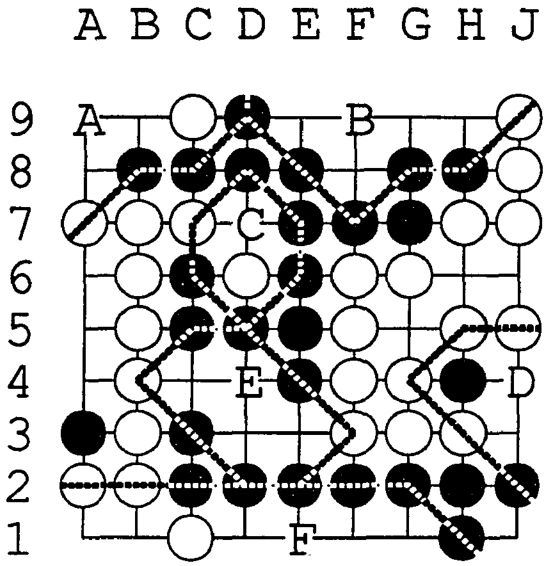
Partitioning the position into its significant tiles.

Figure 4.2 demonstrates the first step to the solution of Berlekamp's problem: partition the whole board into unsettled tiles. Unsettled tiles are areas surrounded by living stones of both colors with empty points inside that are potential territory for at least one player and invadable by the other player. We use term "settled territory" for territory that consists of areas entirely surrounded by one player with empty points inside, Berlekamp does not use this term. In Berlekamp's problem, the prisoner on A3 along with the three empty points A4, A5, and A6 form settled territory for White worth five points. (A3 is one point of territory along with one point for the prisoner on it.) Likewise J1 is one point of settled territory for Black, and the group surrounding G5, H6, and J6 contain three more points of settled territory for White. Settled territories have no significant part in the solution of Berlekamp's problem. Settled territories serve to even the score when the contested areas favor an opponent. Settled territories form a framework for attaching other tiles.

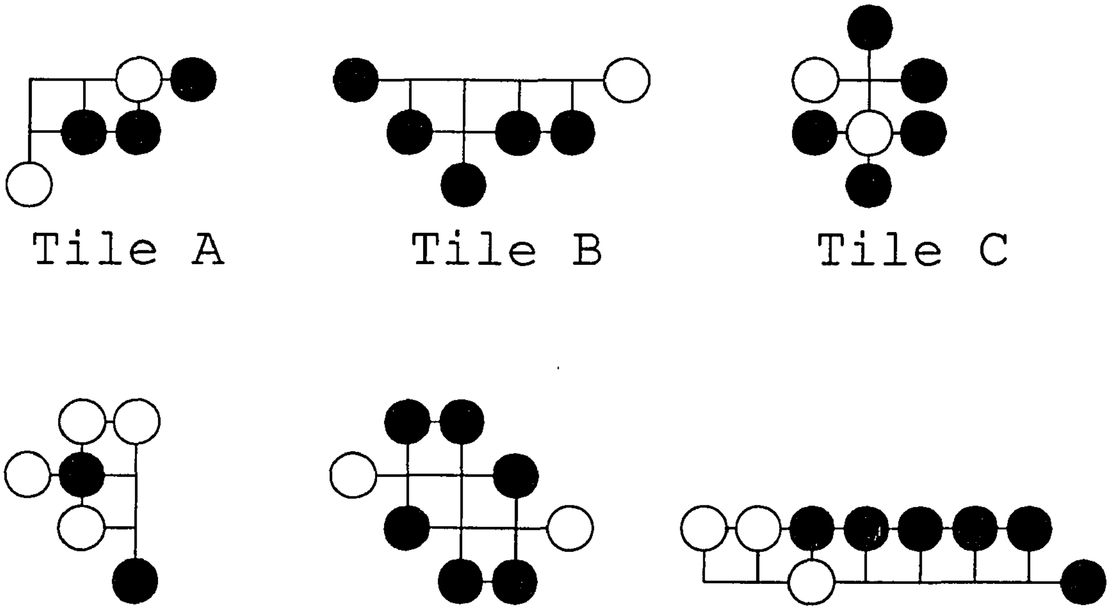
We use dotted lines on the whole board to separate the boundaries of the tiles. When portraying an individual tile, the outer stones are considered immortal, the stones inside the boundaries are potential prisoners (in this example: C9, D6, H4, and C1). To avoid confusion, immortal stones will often be labeled with an "I."

Mathematical Go treats each tile as totally separate. Wolfe's thesis discusses various cases of tiles with minimal interaction but the interaction does not affect the best play within a tile, only the values of a play. This will be explored in Chapter 5.

The partitioning of tiles and the recognition of live groups, is not part of the domain of


A B C D E F G H J
9
8
7
6
5
4
3
2
1
A
B
C
D
E
F

The board
partitioned into
its unsettled tiles.
The first step in the
Mathematical Go
solution.



Tile A Tile B Tile C

Tile D Tile E Tile F

Figure 4.2

Mathematical Go per se. Mathematical Go deals in the evaluation of individual tiles, the comparison of tiles, and the summation of tile values. It also provides for best play game trees within a tile. The partitioning of tiles is done by using ordinary go knowledge. The problem solver and the problem composer must be aware that all large groups are alive.

Having completed the first major step in the Mathematical Go analysis of a go endgame problem, we ask what would have happened if some area of the board were not so clearly partitioned? What if one corner were threatened with life or death? The answer to these questions is that the problem would be a go problem but not a Mathematical Go problem. Mathematical Go solves problems that fall into the domain of problems it can solve. If an area of attack and defense of territory includes a subregion whose life and death status is unclear, then the entire region becomes part of a larger tile, and the application of Mathematical Go to that larger tile may not be accurate. Mathematical Go gives exact evaluations for tiles where the best move in that tile is worth approximately two points or less. Tiles with larger values are beyond its bounds.

Simplify the value of each tile as much as possible

Chapter 2 explained how every finite game can be simplified to a unique canonical form by exhaustively using the operations of deleting dominated moves and bypassing reversible moves. These two operations are used to simplify Mathematical Go positions. Additionally, one can use go knowledge as a heuristic for not looking at bad plays. Specifically, making a move that forms settled territory for an entire area is always worth more than leaving part of an area as unsettled. We call this strategy "settling the position." The simplification of the six tiles in Berlekamp's problem will now be demonstrated.

In Mathematical Go, repeating a position is always allowed. Berlekamp calls games that can be drawn by repetition "loopy." In his paper, Berlekamp remarks on the uniqueness of canonical forms in general [BE2, p. 83]:

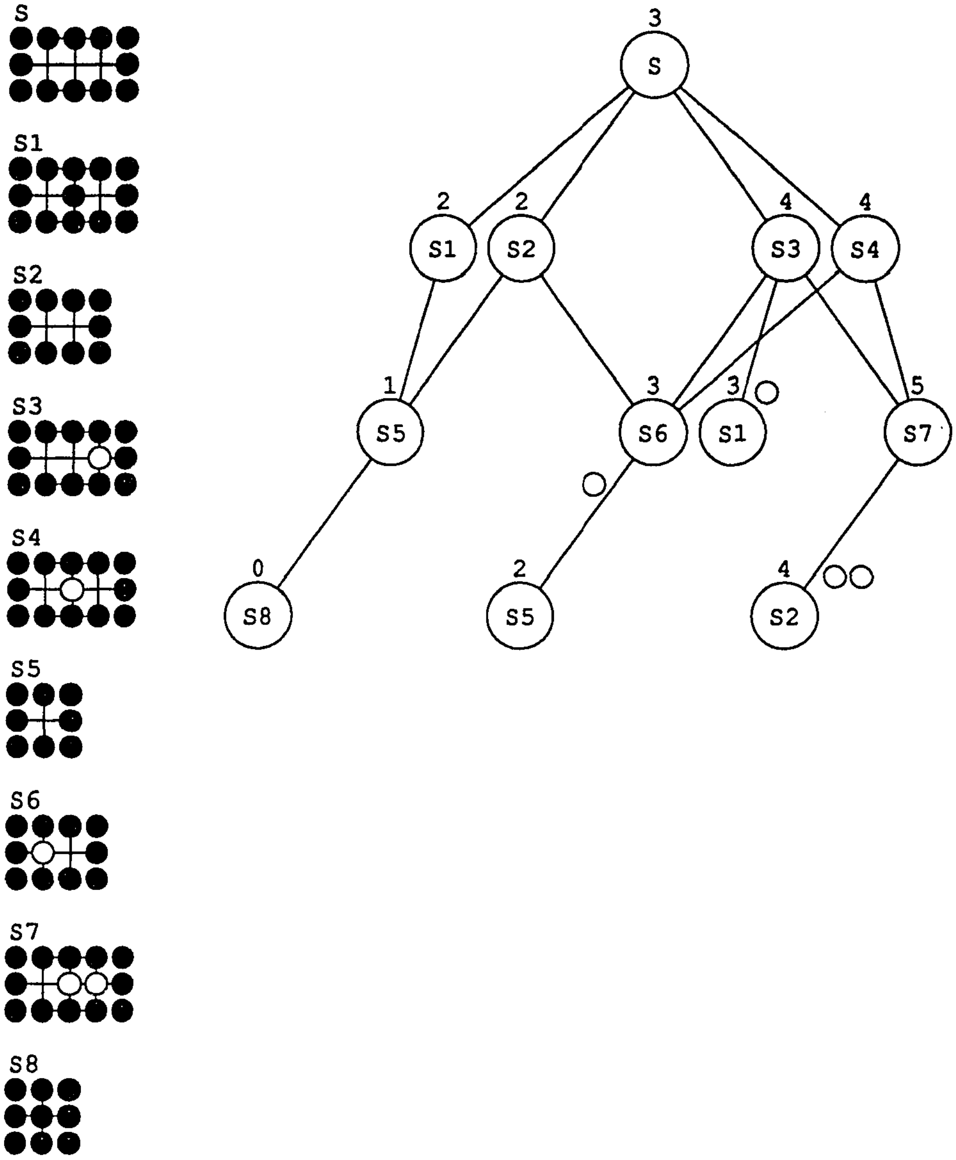
> Potentially unique canonical forms for games with draws are not yet understood. It may be that there are some additional simplification theorems for loopy games, and that the application of such theorems will ensure uniqueness of a canonical form in general. It may also be that the canonical form defined above is already unique: neither a proof nor any counterexample is yet known.

In simplifying games, it is important to remember that $G \geq H$ is verified by $G + (-H) \geq 0$ which means that Black playing second will win on $G - H$; and $G \leq H$ is verified by $G + (-H) \leq 0$ which means that White playing second will win on $G - H$.

Best play game tree

The brute force method for simplifying a tree is to first create the exhaustive game tree and then simplify from the bottom up. We illustrate possible best play from either side by drawing solid lines from node to node. The plays deleted or bypassed are drawn as dashed lines. The actual reversible move and the reversing move are drawn as solid lines and the intermediate nodes are preserved to show the plays made.

Figure 4.3 shows the exhaustive game tree for a settled territory of size three. The actual tiles are displayed on the left. The moves from the original tiles are displayed on the right by


S
S1
S2
S3
S4
S5
S6
S7
S8
3
S
2
S1
2
S2
4
S3
4
S4
1
S5
3
S6
3
S1
5
S7
0
S8
2
S5
4
S2

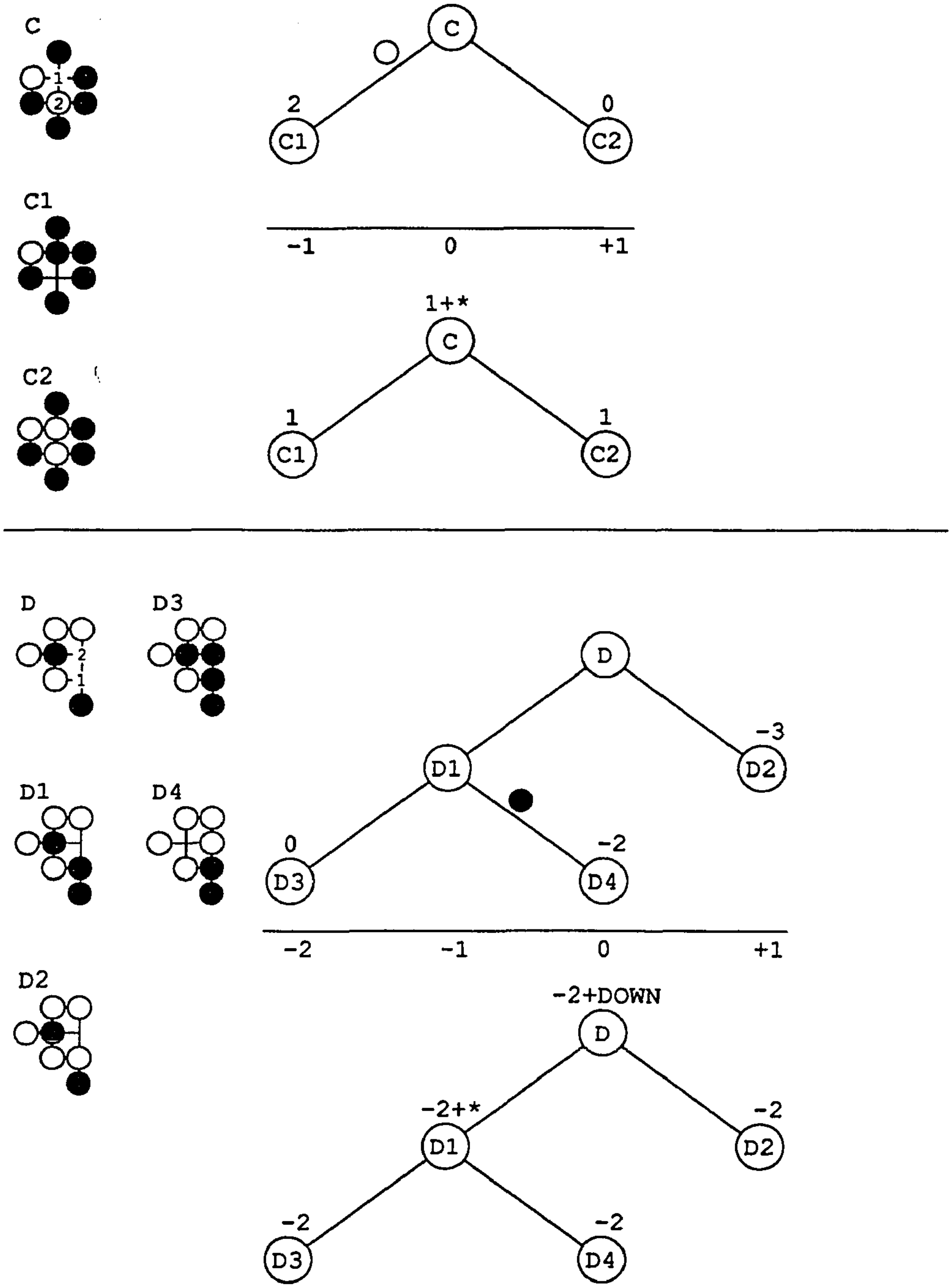
The exhaustive WW game tree for settled territory of size three.

Figure 4.3

arcs connecting the named tiles. Prisoners captured on the move are shown by White (or Black as the case may be) stones alongside the arc. The thing to note about settled territory is that there is no strategy at all in the play, merely filling up empty points. Best play and worst play yield identical results. In game trees the value for each leaf node (nodes from which there are no moves) is shown. Duplicate nodes are not repeated (in our example, S1, S2, and S5 are only exploded once). We will adopt the convention that if a node is settled territory then it will not be exploded. The value of a leaf node is the value of the territory surrounded plus the value of any prisoner inside the territory, plus the net sum of all prisoners captured in play from the root to the leaf node. We shall see below that the chilled value of settled territory is the same as the value before chilling, so there is another reason for not exploding these nodes.

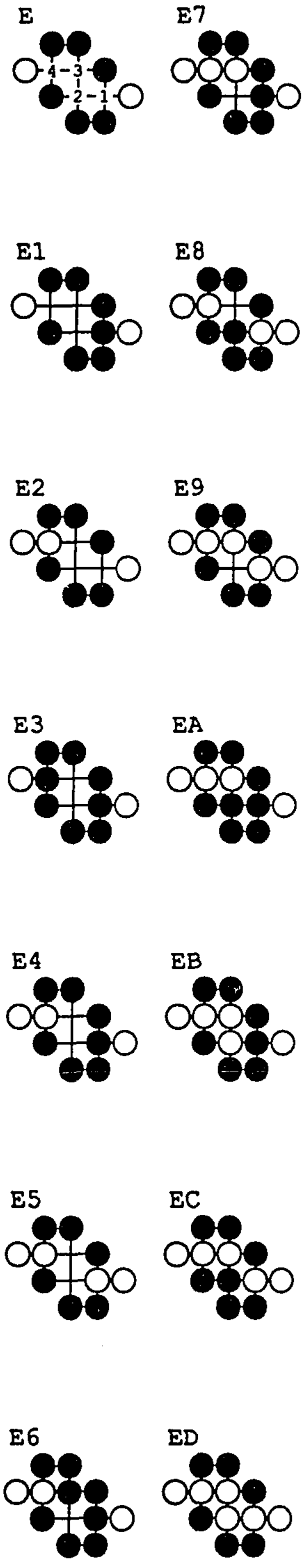
Figure 4.4 shows the best play game tree for Tile C and Tile D of Berlekamp's problem. In this and the next several figures, the best play game trees are shown above the dividing line labeled with signed numbers. The "chilled" trees below will be discussed under step III. As mentioned above, Tiles C1, D2 and D4 are leaf nodes because the territory is settled, even though moves are available for play. In Tile D, the move to D2, point 1 of Tile D, settles the position and dominates any other White play. As suggested by the go proverb, "the opponent's best move is your own," a Black play at 1 dominates any other Black play. It is a simple exercise to constructively compare Black's move to point 1 versus the negative of the game after Black's move to point 2 as a win for Black going second.

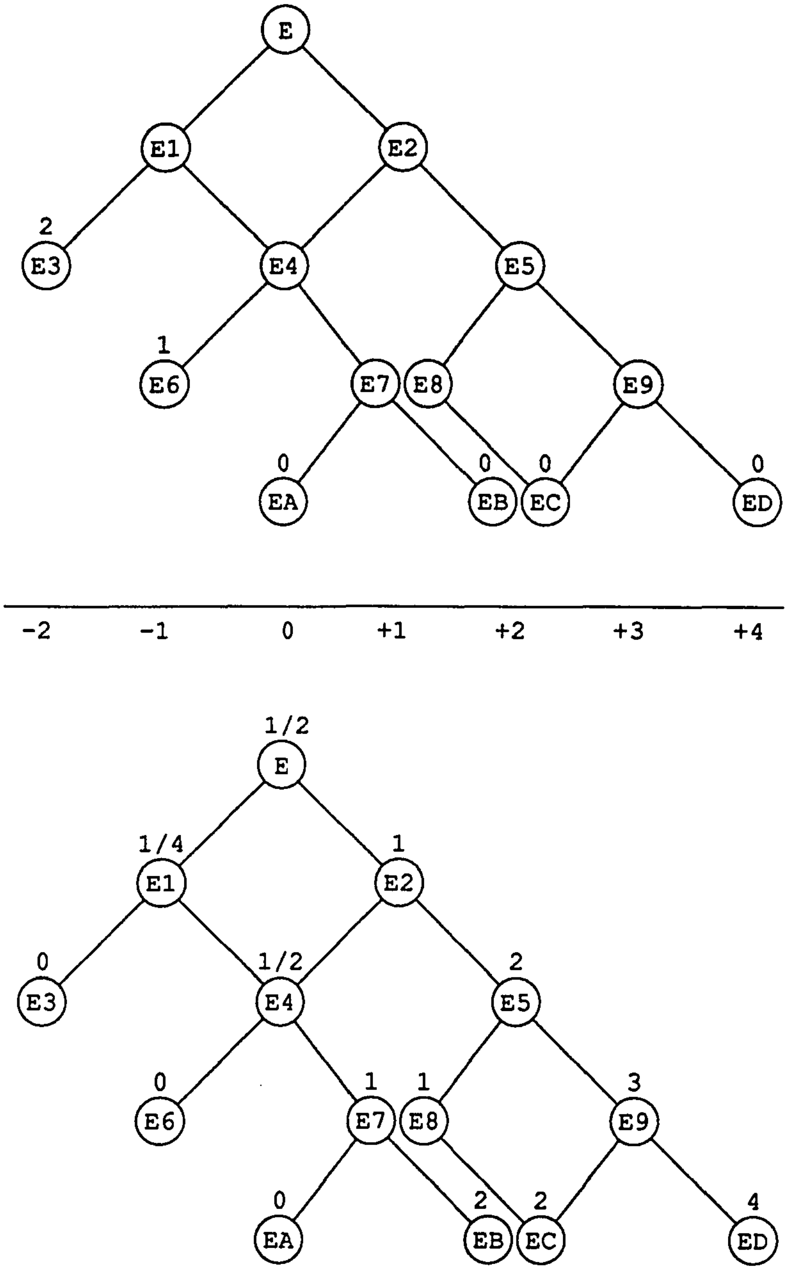
Figure 4.5 shows the best play game tree for Tile E of Berlekamp's problem. Plays to points 1 or 4 of Tile E dominate plays to points 2 or 3. Symmetry reduces the tree by half. The leaf nodes EA, EB, EC, and ED are shown as separate only to indicate the actual play.


C
1
2
C
2
0
C1
C2
C1
-1
0
+1
1+*
C
C2
1
1
C1
C2
D
2
1
D3
D
-3
D1
D2
D1
D4
0
-2
D3
D4
-2
-1
0
+1
D2
-2+DOWN
D
-2+*
-2
D1
D2
-2
-2
D3
D4

Upper half of diagram: best play game tree and chilled tree morphism for Tile C of the whole board problem.
Lower half of diagram: best play game tree and chilled tree morphism for Tile D of the whole board problem.

Figure 4.4


E
4
3
2
1
E7
E1
E8
E2
E9
E3
EA
E4
EB
E5
EC
E6
ED


E
E1
E2
2
E3
E4
E5
1
E6
E7
E8
E9
0
EA
0
EB
0
EC
0
ED
-2
-1
0
+1
+2
+3
+4
1/2
E
1/4
E1
1
E2
0
E3
1/2
E4
2
E5
0
E6
1
E7
1
E8
3
E9
0
EA
2
EB
2
EC
4
ED

Best play game tree and
chilled tree morphism for
Tile E of the whole board problem.

Figure 4.5

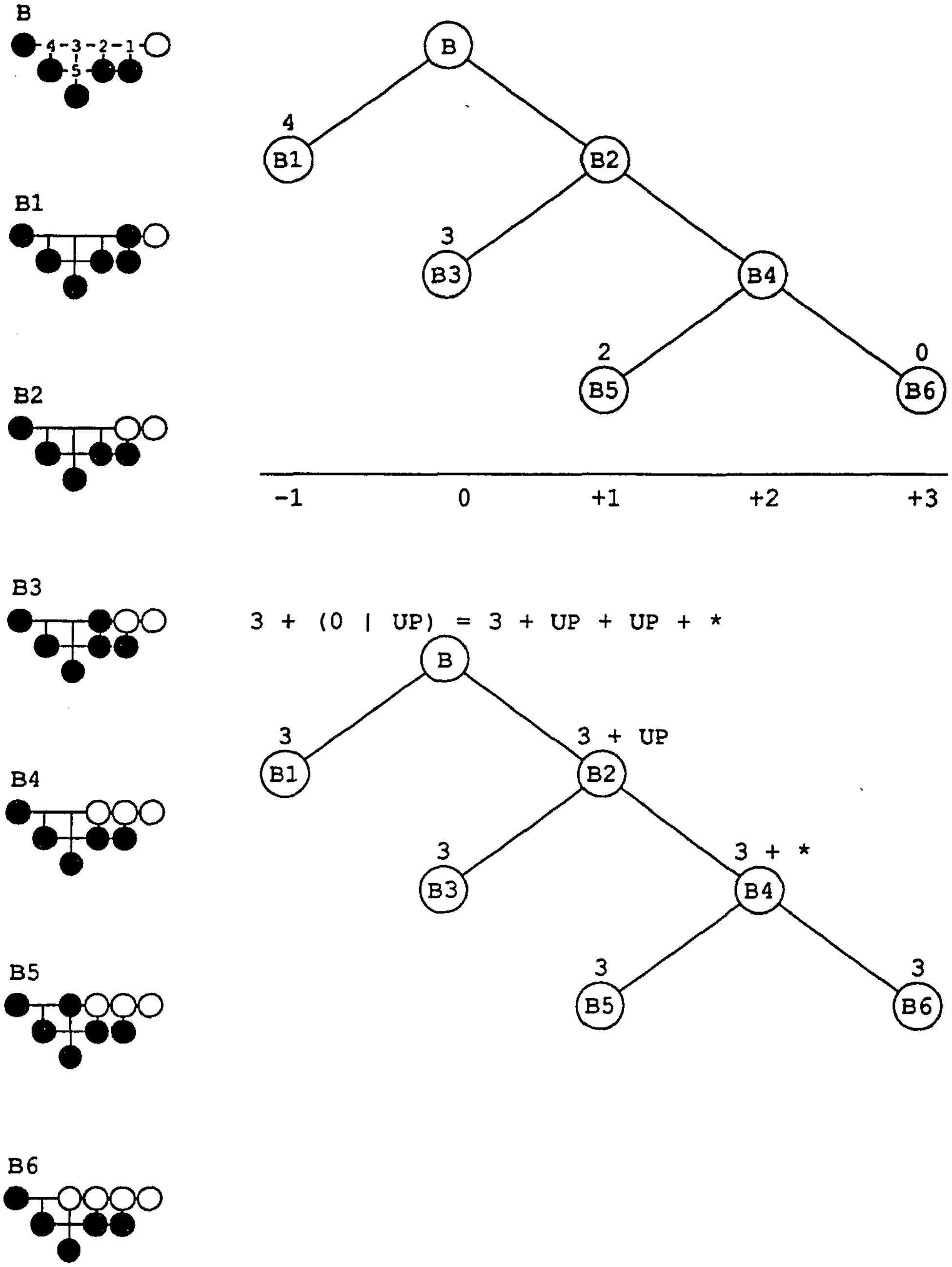
All tiles with no play are essentially identical, with value zero. Notice that E4 can be reached by two paths E.

Figure 4.6 shows the best play game tree for Tile B of Berlekamp's Problem. Plays that immediately secure the most territory dominate other Black plays, and , once again, White does best to follow the proverb, "the opponent's best move is your own." Exploding Tile B6 would show the outcome to be a second player win, hence it has a value of 0.

Figure 4.7 shows the best play game tree for Tile F of Berlekamp's problem. The exhaustive game tree has over fifty tiles. The Black moves from F2, F6, and FA are reversible. The next figure also has reversible moves with a much smaller exhaustive tree, and reversible moves will be discussed there.

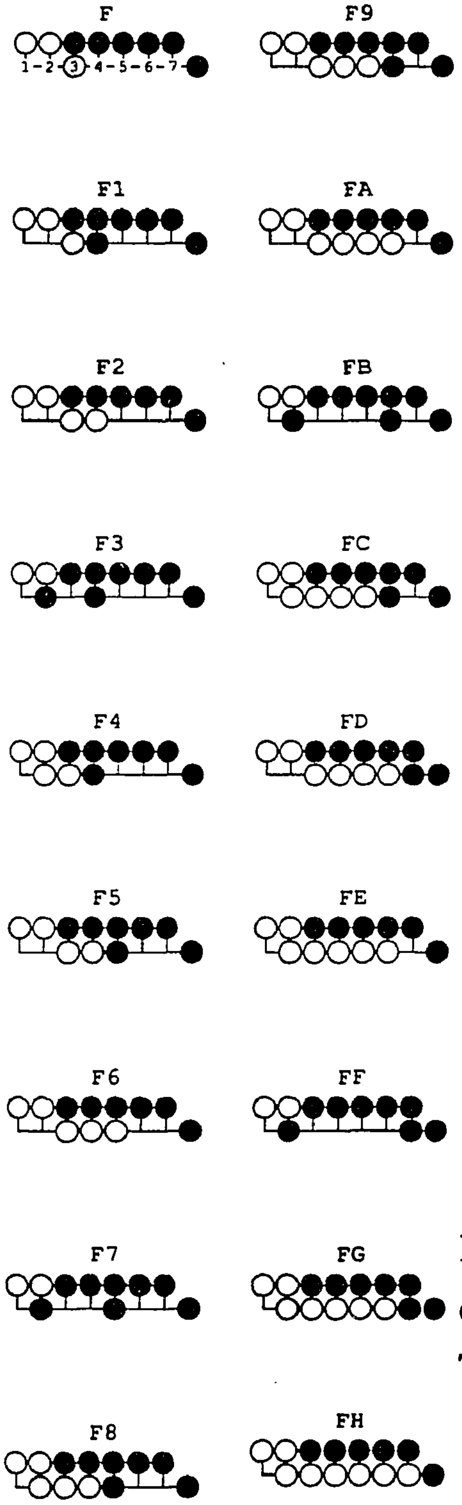
Figure 4.8 shows the nearly complete unsimplified game tree for Tile A of Berlekamp's problem. Figure 4.9 shows the best play game tree for the same tile. Figure 4.8 includes the WW values for all nodes only to verify subsequent simplification. In Figure 4.8 Tile A5 appears in the path A5, A9, AB, A5, a cycle of length three. Tile AF and Tile AD appear in the path AF, AD, AF, AD, a cycle of length two from either tile. These two tiles illustrate ko. This is a true loopy position.

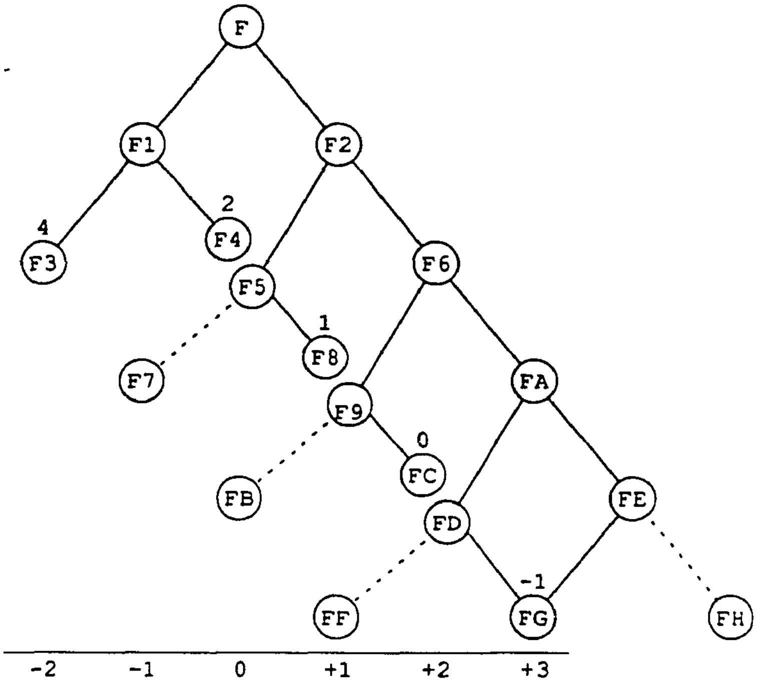
The move from AB to AF is reversible by a move to AE. AE is as least as good for White as AB. Likewise the move from AB to AC is reversible by a move to AE. The diamond shaped pattern formed by AB, AF, AE, AC is easily recognizable in any game tree as the game 0, a second player win. In expanding a node, when one sees this zero game, one may ignore drawing the edges to bypassed nodes such as AD and AA. (The notation 1 above AB

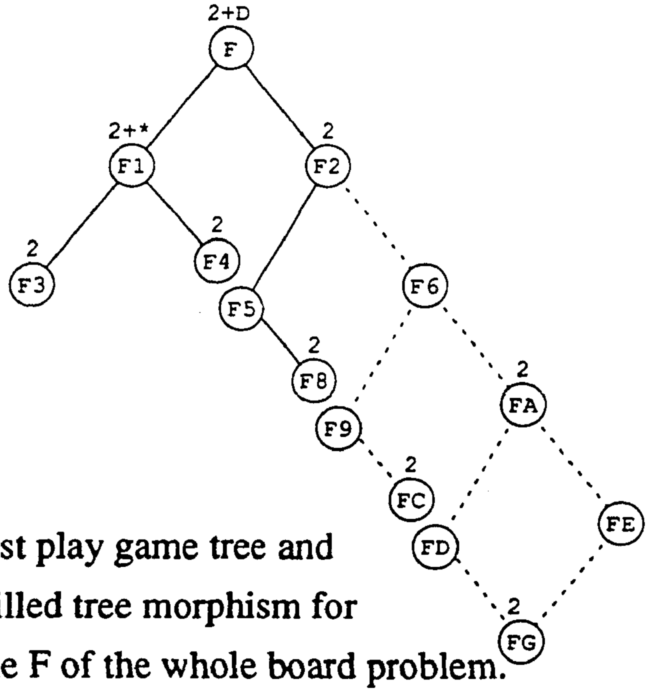

B
4-3-2-1
5
B1
B2
B3
B4
B5
B6
B
4
B1
B2
3
B3
B4
2
B5
0
B6
-1
0
+1
+2
+3
3 + (0 | UP) = 3 + UP + UP + *
B
3
B1
3 + UP
B2
3
B3
3 + *
B4
3
B5
3
B6

Best play game tree and
chilled tree morphism for
Tile B of the whole board problem.

Figure 4.6

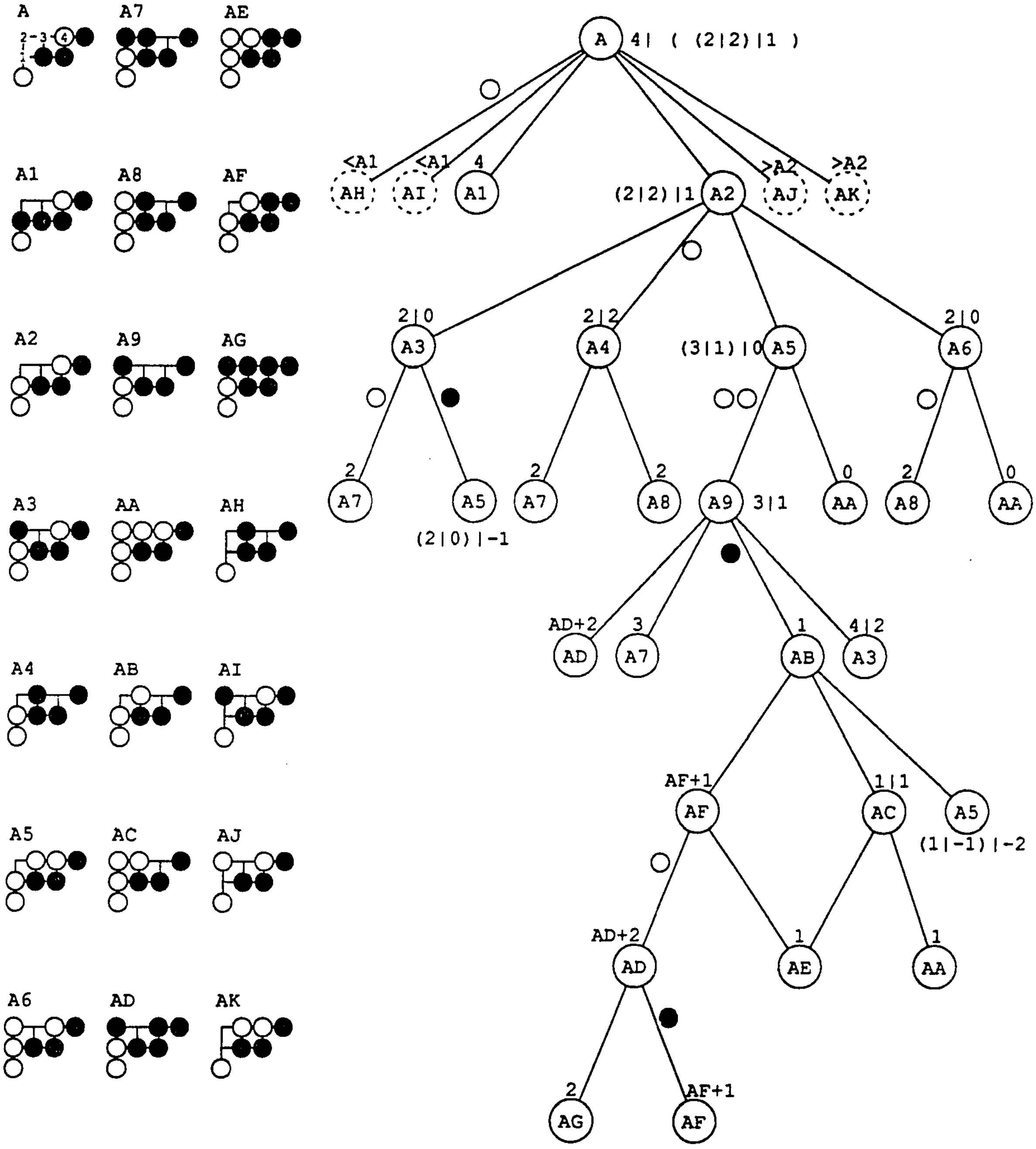

F
1-2-3-4-5-6-7
F9
F1
FA
F2
FB
F3
FC
F4
FD
F5
FE
F6
FF
F7
FG
F8
FH


F
F1
F2
4
F3
2
F4
F6
F5
1
F7
F8
FA
F9
0
FB
FC
FE
FD
-1
FF
FG
FH
-2
-1
0
+1
+2
+3

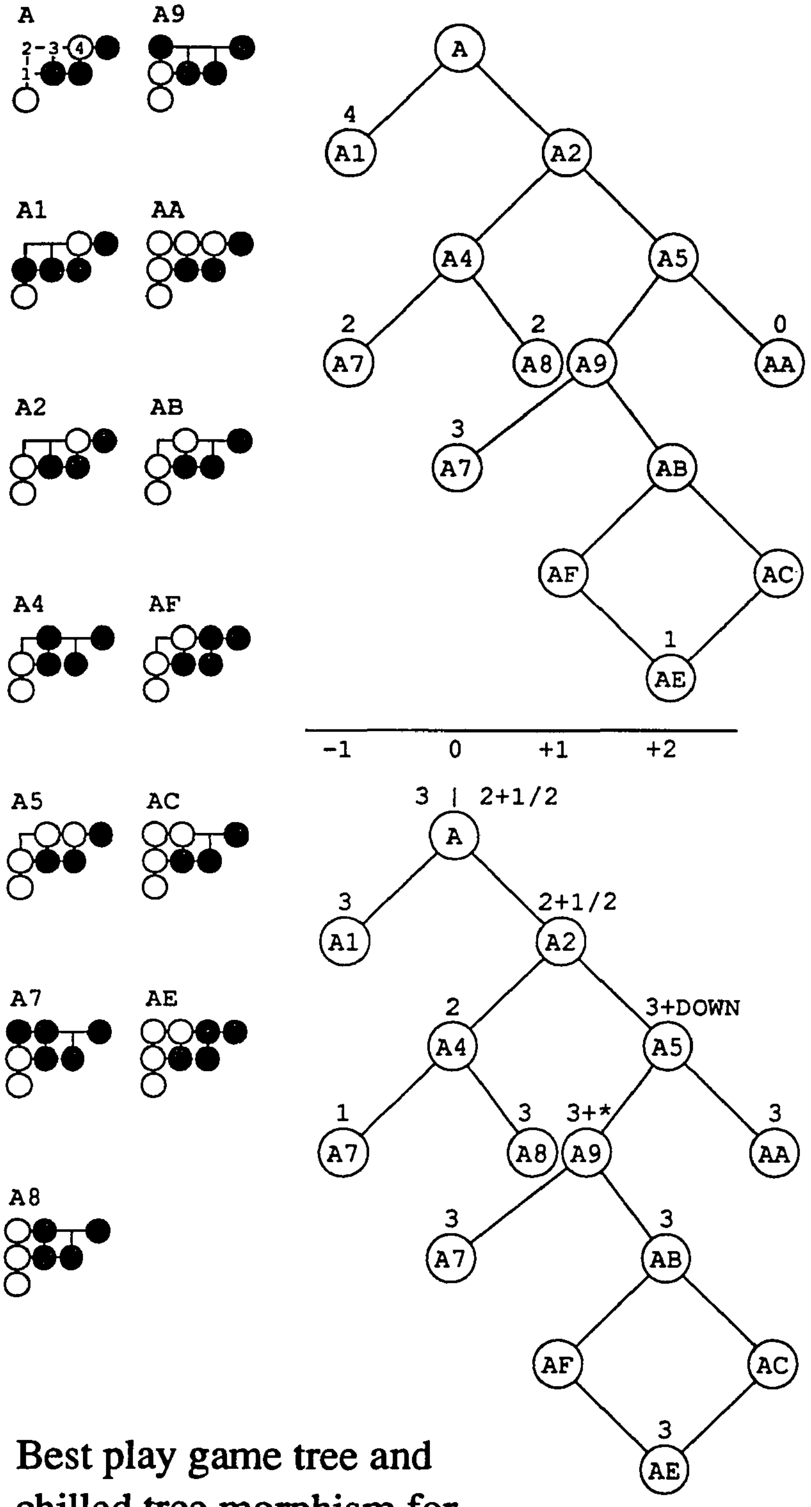

2+D
F
2+*
F1
2
F2
2
F3
2
F4
F6
F5
2
F8
2
FA
F9
2
FC
FE
FD
2
FG

Best play game tree and chilled tree morphism for Tile F of the whole board problem.

Figure 4.7


A
A7
AE
2-3-4
1
A1
A8
AF
A2
A9
AG
A3
AA
AH
A4
AB
AI
A5
AC
AJ
A6
AD
AK
A 4| ((2|2)|1)
<A1
<A1
4
>A2
>A2
AH
AI
A1
(2|2)|1 A2
AJ
AK
2|0
2|2
(3|1)|0 A5
2|0
A3
A4
A6
2
A7
A5
(2|0)|-1
2
A7
2
A8
A9 3|1
0
AA
2
A8
0
AA
AD+2
AD
3
A7
1
AB
4|2
A3
AF+1
AF
1|1
AC
A5
(1|-1)|-2
AD+2
AD
1
AE
1
AA
2
AG
AF+1
AF

The nearly complete unsimplified game tree for
Tile A of the whole board problem
Figure 4.8


A
2-3-4
1
A9
A1
AA
A2
AB
A4
AF
A5
AC
A7
AE
A8
A
4
A1
A2
A4
A5
2
A7
2
A8
A9
0
AA
3
A7
AB
AF
AC
1
AE
-1
0
+1
+2
3 | 2+1/2
A
3
A1
2+1/2
A2
2
A4
3+DOWN
A5
1
A7
3
A8
3+*
A9
3
AA
3
A7
3
AB
AF
AC
3
AE

Best play game tree and
chilled tree morphism for
Tile A of the whole board problem

Figure 4.9

refers to the Black prisoner captured in playing from A.) The move from A9 to A3 is dominated by the play to AB. Using the simplified value of A5 after A9 has been simplified, A4 now dominates A3. Again using the simplified form for A5, it now dominates A6 from A2.

After all this simplification, no loops remain in the best play game tree. Each player can play for a larger value than would have been obtained by aiming for a draw by repetition. This is the key to the analysis of many potentially "loopy" Mathematical Go positions. In Chapter 8 we look at counter examples.

Chill the game

Chilling a game is a mathematical operation on a WW game (or game tree, or tile) that maps the original set of values to a new set of values. Berlekamp and Wolfe use the same board for the regular game and the chilled game. Berlekamp implements "markings," small black and white circles on the board to indicate the chilled value. Berlekamp and Wolfe both note that markings can be confusing. Markings are only a notational convenience. To avoid confusion, instead of markings, we will use separate game trees for the WW values calculated in step II and the chilled game trees calculated now in step III.

Recalling our discussion at the end of Chapter 2, a game that neither player has an incentive to play in has a value that is a number and is called cold. A game where the players do have an incentive to play in is called hot. Cooling maps hot games to cold games, chilling is a form of cooling.

Cooling a game:

G_t (G cooled by t) is defined for increasing values of t as follows:

$G_t = \{ G^B{}_t, -t \mid G^W{}_t, +t\}$

unless there is a $w < t$ for which G_w is infinitesimally close to a number x, in which case

$G_t = x$ for all $t > w$.

Cooling a hot game G by t means: if a player makes a move in a hot component (which he will if there is one) then he must pay a premium of t for that move unless the result would be less than a number, in which case the resulting cooled value is that number. The smallest number t_0, such that G cooled by t_0 is a number m, is called the *temperature* of G. m is called the *mean value* of G, m(G). For any hot game G, if ε is any positive number greater than 0, we have:

$$m(G) - t_0 - \varepsilon < G < m(G) + t_0 + \varepsilon.$$

Chilling is cooling by temperature 1. Tiles whose values are already numbers, whether integer or fraction, won't have their values changed by chilling. Tiles with hot values don't have numeric values but they do have mean values. Chilling reduces the mean value of a tile by at most 1, and no more than the original mean value. As we shall see, chilling tiles may yield values that will give specific strategic insights for best play.

To correctly annotate the chilled game tree with the correct values, start with the best play game tree using the values for the leaf nodes. For each of the game trees we show a chilling scale which conveniently sums net moves made by each player. For each leaf node in the best play game tree, sum its value with the value of the chilling scale directly below it, and place the result in the chilled tree. Then, percolate the value of the leaf nodes upwards to the root, using the formulas for evaluating games that were discussed in Chapter 2.

Let us review the chilled game trees for the six tiles in Berlekamp's problem. In Figure 4.4, for Tile C, we see that the game $\{2 \mid 0\}$ chills to the game $\{1 \mid 1\}$. WW likes to represent games in a concise notation. First, we can always factor out a number from each side, in this case 1.

$$\{1 \mid 1\} = 1 + \{0 \mid 0\} = 1 + *.$$

Next, the sum of two games is usually represented by the catenation of two games. In WW notation, the symbol "+" is used to represent a class of values known as "tiny." We won't use "+" to represent "tiny," so to avoid confusion, we leave "+" to indicate sum in the usual way. In general

$$\{n \mid n\} = n + \{0 \mid 0\} = n + *$$

which in WW terms is n*, where n is a number.

In Figure 4.4 for Tile D, the game $\{-2 \mid -2\} = -2 + *$, as explained above. There are two small games which occur so often, they have there own names

$$\{0 \mid 0 \mid 0\} = \{0 \mid * \} = \text{UP}.$$

$$\{0 \mid 0 \mid 0\} = \{* \mid 0 \} = \text{DOWN}.$$

Factoring the number out of $\{-2+* \mid -2\}$ becomes -2 + DOWN.

In Figure 4.5, all internal nodes have Black moves smaller than their corresponding White moves, hence every node is a number, the simplest number that fits between both options (as discussed in Chapter 2). This tile illustrates that chilled Mathematical Go positions can have *fractional* values.

In Figure 4.6, the node for B2 is simply the negative tree of node D in Figure 4.4. The simplifying equalities

$\{0 \mid UP\} = UP + UP + *$ and

$\{DOWN \mid 0\} = DOWN + DOWN + *$

can be demonstrated in the usual WW fashion of taking the negative of one side and demonstrating a second player win making the value of the difference 0. These two simplifications occur often in Mathematical Go tiles.

In Figure 4.7, much simplification has occurred, either as a result of bypassing a reversible move, or simplifying an internal node to 0. When a path such as F2, F5, F8 is shown, it means that the move from F5 to F7 has been bypassed. If this were a WW simplified game tree, F2 would be a leaf node with value 2, but we are preserving the plays until the actual settled territory is reached. In this case, the value of 2 is brought upward from F8 to F2. Ignoring the number and looking at the hot part of the game, Tile F is identical with Tile D, each having the tree for DOWN.

In Figure 4.9, both AF and AC were reversing moves. Again, if this were a WW game tree, AB would be a leaf node with value 3, but we are preserving the path to AE which is settled territory. The rest of the game summing follows the rules mentioned above. As we shall explain in the next paragraph, Tile A is interesting in that it is the only game in this tile set with an incentive greater than an infinitesimal.

Play out the game always selecting the largest incentive

We will now give some rules for how to use the chilled values of game tiles to determine optimal play. In the game G = {B | W}, Black's incentive is defined as B - G, and White's incentive is defined as G -W. In other words, a player's incentive to play in a particular tile is the difference in value between the original game, and the game after a move is played. Since values are additive, and a game is the disjunctive sum of its constituent tiles, this means we should always play in the tile with the largest incentive.

Hot games fall into two classes, games with finite incentives and games with infinitesimal incentives. For example, the hot game {4 | 2} has a temperature of 1, a mean value of 3, an alternative representation of value as 3 + {1 | -1}, and an incentive of 1 for both players. The game {5 | 5} has a temperature of 0, a mean value of five, an actual value of 5+*, and an incentive of * for both players. In a hot game with finite value, whoever plays first will receive more value, which will translate into more moves. In a hot game with infinitesimal value, the player who moves first will benefit relating to who gets the last move, but there will be no extra bankable moves.

There is a partial ordering among infinitesimals. * is the largest infinitesimal, it is neither positive nor negative. UP is the largest positive infinitesimal, DOWN is the largest negative infinitesimal. "Tiny" is the smallest class of positive infinitesimals, the game tiny-n is defined by $tiny_n$ = { 0 | 0 | n}. "Miny" is the smallest class of negative infinitesimals, the game miny-n is defined by $miny_n$ = { n | 0 | 0}. WW lists all the simplest forms for UPs, DOWNs, and STARs [BE1, p. 73]. Wolfe lists the partial order of most infinitesimal incentives for tiles that may arise in corridors [WD1, p.35]. (We will discuss corridors in Chapter 5.) When a new comparison has to be made, one that isn't already listed in the above references, we can always make the constructive comparison as discussed in Chapter 2. For the problem under consideration, we only need to know that UPs and DOWNs are additive, the more UPs in a tile the better for Black: the more DOWNs the better for White.

Figure 4.10 describes White's win against stern resistance from Black. The figure is organized as follows. Before each move, the current set of tiles is listed, along with the possible moves by Black and White. The names of the tiles correspond to the names of the tiles in Figures 4.4, 4.5, 4.6, 4.7, and 4.9. Alongside the name of each tile is its value. On the right of each tile, the move made is indicated. If any other move has equal value, it is so listed. When a tile is settled territory, the Black and White moves are no longer listed.

Move 1: A is a hot game with finite value. The original game has a mean value of $2\frac{3}{4}$. By playing here, White increases his value by $\frac{1}{4}$.

Move 2: Playing in B would use up a valuable infinitesimal. In a choice of infinitesimals, as in a choice among numbers, one plays in the smallest value which yields the least decrement of value. Here playing in either DOWN has equal value. Playing in B gets rid of UP+UP+STAR (see above identity).

Move 3: Notice that the incentives to play in B were asymmetrical. B to B2 is greater than F to F2.

Move 4: Black wishes to remove White's DOWNs while saving his UPs. The converse goal applies for White.

Move 5: Removing an UP by White is the only choice.

Moves 6-9: Playing in STARs before numbers.

Move 10: All tiles are numbers. A play in E loses only $\frac{1}{4}$.

Figure 4.10 page 1 of 3.

```
prior to move: 1                      |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A  (A1|A2) | A1 3       | A2 2+1/2   | MOVE MADE
 B 3+(0|UP) | B1 3       | B2 3+UP    |
 C      1+* | C1 1       | C2  1      |
 D  -2+DOWN | D1 -2+*    | D2 -2      |
 E      1/2 | E1 1/4     | E2  1      |
 F   2+DOWN | F1 2+*     | F2  2      |
--------------------------------------+
prior to move: 2                      |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4 2       | A5  3+DOWN |
 B 3+(0|UP) | B1 3       | B2 3+UP    |
 C      1+* | C1 1       | C2  1      |
 D  -2+DOWN | D1 -2+*    | D2 -2      | MOVE MADE
 E      1/2 | E1 1/4     | E2  1      |
 F   2+DOWN | F1 2+*     | F2  2      | EQUAL CHOICE
--------------------------------------+
prior to move: 3                      |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4 2       | A5  3+DOWN |
 B 3+(0|UP) | B1 3       | B2 3+UP    | MOVE MADE
 C      1+* | C1 1       | C2  1      |
 D1     -2+* | D3 -2      | D4 -2      |
 E      1/2 | E1 1/4     | E2  1      |
 F   2+DOWN | F1 2+*     | F2  2      |
--------------------------------------+
prior to move: 4                      |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4 2       | A5  3+DOWN |
 B2    3+UP | B3 3       | B4 3+*     |
 C      1+* | C1 1       | C2  1      |
 D1    -2+* | D3 -2      | D4 -2      |
 E      1/2 | E1 1/4     | E2  1      |
 F   2+DOWN | F1 2+*     | F2  2      | MOVE MADE
--- ----------------------------------+
prior to move: 5                      |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4 2       | A5  3+DOWN |
 B2    3+UP | B3 3       | B4 3+*     | MOVE MADE
 C      1+* | C1 1       | C2  1      |
 D1    -2+* | D3 -2      | D4 -2      |
 E      1/2 | E1 1/4     | E2  1      |
 F1     2+* | F3 2       | F4  2      |
--------------------------------------+
prior to move: 6                      |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4  2      | A5  3+DOWN |
 B4     3+* | B5  3      | B6  3      | MOVE MADE
 C      1+* | C1  1      | C2  1      | EQUAL CHOICE
 D1    -2+* | D3 -2      | D4 -2      | EQUAL CHOICE
 E      1/2 | E1  1/4    | E2  1      |
 F1     2+* | F3  2      | F4  2      | EQUAL CHOICE
--------------------------------------+
```

Figure 4.10 page 2 of 3.

```
prior to move: 7                     |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4   2     | A5  3+DOWN |
 B5       3 |            |            |
 C      1+* | C1   1     | C2   1     | EQUAL CHOICE
 D1    -2+* | D3  -2     | D4  -2     | EQUAL CHOICE
 E      1/2 | E1   1/4   | E2   1     |
 F1     2+* | F3   2     | F4   2     | MOVE MADE
-------------------------------------+
prior to move: 8                     |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4   2     | A5  3+DOWN |
 B5       3 |            |            |
 C      1+* | C1   1     | C2   1     | EQUAL CHOICE
 D1    -2+* | D3  -2     | D4  -2     | MOVE MADE
 E      1/2 | E1   1/4   | E2   1     |
 F4       2 |            |            |
---- --------------------------------+
prior to move: 9                     |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4   2     | A5  3+DOWN |
 B5       3 |            |            |
 C      1+* | C1   1     | C2   1     | MOVE MADE
 D3      -2 |            |            |
 E      1/2 | E1   1/4   | E2   1     |
 F4       2 |            |            |
-------------------------------------+
prior to move: 10                    |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4   2     | A5  3+DOWN |
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E      1/2 | E1   1/4   | E2   1     | MOVE MADE
 F4       2 |            |            |
-------------------------------------+
prior to move: 11                    |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4   2     | A5  3+DOWN |
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E1     1/4 | E3   0     | E4   1/2   | MOVE MADE
 F4       2 |            |            |
-------------------------------------+
prior to move: 12                    |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4   2     | A5  3+DOWN | EQUAL CHOICE
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E4     1/2 | E6   0     | E7   1     | MOVE MADE
 F4       2 |            |            |
-------------------------------------+
```

Figure 4.10 page 3 of 3.

```
prior to move: 13                       |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A2   2+1/2 | A4  2      | A5  3+DOWN |  MOVE MADE
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E6       0 |            |            |
 F4       2 |            |            |
--------------------------------------+
prior to move: 14                       |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A5  3+DOWN | A9  3+*    | AA  3      |  MOVE MADE
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E6       0 |            |            |
 F4       2 |            |            |
--------------------------------------+
prior to move: 15                       |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 A9     3+* | A7  3      | AB  3      |  MOVE MADE
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E6       0 |            |            |
 F4       2 |            |            |
--------------------------------------+
prior to move: 16                       |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 AB       3 | AF --> 3   | AC --> 3   |  MOVE MADE
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E6       0 |            |            |
 F4       2 |            |            |
--------------------------------------+
prior to move: 17                       |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 AC       3 | AE 3       | AE 3       |  MOVE MADE
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E6       0 |            |            |
 F4       2 |            |            |
--------------------------------------+
after    move: 17                       |
CURRENT TILE|BLACK'S MOVE|WHITE'S MOVE|
 AE       3 |            |            |
 B5       3 |            |            |
 C2       1 |            |            |
 D3      -2 |            |            |
 E6       0 |            |            |
 F4       2 |            |            |
--------------------------------------+
```

Moves 11-13: More plays in the least disincentive.

Moves 14-17: At this point, only the resultant tiles from the original Tile A are in play. It is important to remember that the best play game tree kept those nodes which would have been deleted from the canonical form game tree. A move to AB at move 15 is really a move to AE. The tile AF is bypassed during simplification.

This chapter has provided us with steps to determine optimal play for a specific problem that can be applied to any general problem of this type. In the next chapter we will look at Wolfe's added measure of complexity to Mathematical Go, sockets. We will also look at Wolfe's treatment of certain shaped tiles, which he calls corridors.

CHAPTER 5. Corridors and sockets

This chapter will discuss some of Wolfe's [WD1] results. Let's review what was known before Wolfe's work. In Chapter 2, we saw games defined recursively as an ordered pair of sets of games. Numbers were defined by a relation on that ordered pair, in the fashion of Dedekind. Thus, numbers, or cold games, were games with the number property. This representation of games allows for many extensions. In Chapter 2, we saw that this representation over finite games generates all positive and negative dyadic fractions, using the two operations of disjunctive sum and negation, (i.e., addition and subtraction). Conway [CJ] extended this theory outward in the domain of numbers. He described the operations of multiplication and division. He demonstrated that numbers that are cold games form a field, and that subclasses of this field form another field. By looking at infinite numbers that are cold games, he generated all real numbers. This alternate representation of real numbers by ordered sets corresponding to the resolution of two player conflict is not at all intuitive.

In WW [BE1], Berlekamp, Conway, and Guy elucidated game theory by applying their methods to many instances of human playable games. Testing the efficiency of WW theorist versus game-playing domain expert was impractical because these games, though intellectually challenging, were not popular enough to have acquired a recognized theory of play. Among many other things, what Berlekamp [BE2] did was to find a game that provided a level playing field for a test between a WW theorist and a domain specific expert. Go professionals who study for years have vast memories and excellent shape recognition skills, but they use evaluation methods that we think relate to the unchilled game. Berlekamp saw the relation between the chilled game and the true outcome. Playing the chilled game tree, as we have shown in Chapter 4, is much less taxing on the mind than looking at exhaustive game trees.

Berlekamp's work was a breakthrough in psychology as well as in mathematics.

Berlekamp has claimed to have looked at many Mathematical Go dialects, but has published only the one example we discussed in the previous chapter. This particular example within this dialect has several noteworthy features. The tree for Tile A was quite complicated and had several loops. The tiles were totally partitioned. Wolfe's work [WD1] took these two features as points of resistance and moved in different directions. Among other things, Wolfe investigated corridors (i.e., tiles with simply defined values) and sockets (i.e., loosely connected tiles).

Without going into unnecessary technicalities, we note that much of Wolfe's work uses the concept of warming by temperature one, which is the inverse operation of chilling. Recall that chilling is defined as cooling a game by temperature one. Warming is the inverse of cooling. There is no term for specifically warming by one, so Wolfe uses the general term "warming" to mean the specific term "warming by one." Berlekamp, Conway, Guy, and Wolfe,([JC], [BE1], [BE2]. and [WD1]), represent warming with the familiar integral symbol, although warming has nothing to do with either integration or calculus.

The definition of warming for the game $G = \{ G^B \mid G^W \}$ is:

$\int G = G$ if G is an even integer.

$\int G = G + *$ if G is an odd integer.

$\int G = \{ 1 + \int G^B \mid -1 + \int G^W \}$ otherwise.

Berlekamp labels just one of his results in his paper [BE2, p.95] a "BIG THEOREM."

BIG THEOREM: For "normal" go endgames (whose canonical forms are loop-free), chilling can be inverted by warming and then adding 0 or *, according to an appropriate parity.

Wolfe proves this result in his thesis [WD1, pp. 18-21]. What is big about the BIG THEOREM is the equivalence of go and Mathematical Go. If we look at games where the legal play rules of both games intersect (i.e., the entire game tree legal plays are identical), then the outcome under either set of evaluation rules is identical. Chapter 6 will elaborate on this identity in scoring. The intersection of rule sets, the "'normal' go endgame," bars ko, any other draw by repetition, and seki.

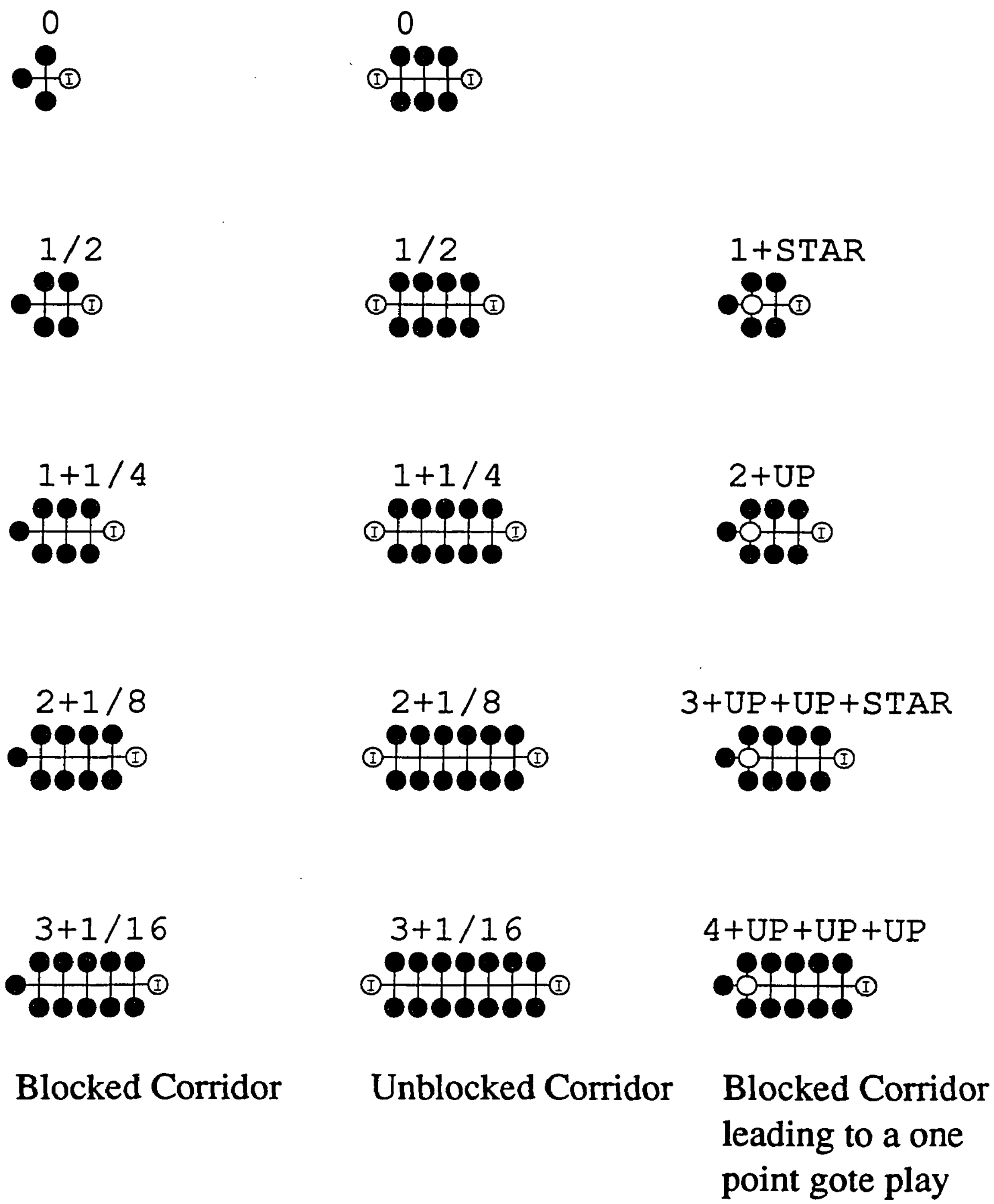
Corridors are a subset of normal go endgames. Corridors have values that can be be calculated recursively, strictly on the basis of class and length. Corridors are tiles with a playing field of width one, and indeterminate length, between either two walls of immortal stones of the same color, or one wall of immortal stones facing the edge of the board. Either one or both ends of a corridor may be reduced by an immortal group of the opponent.

Figure 5.1 illustrates three simple classes of corridors. Blocked corridors, illustrated in the left column, have one entrance point for the invader. Their chilled value is

$$\text{length} - 2 + \frac{1}{2^{length-1}} .$$

Unblocked corridors, illustrated in the middle column, have two entrance points for the invader. Their chilled value is

$$\text{length} - 4 + \frac{1}{2^{length-3}} .$$


0
0
1/2
1/2
1+STAR
1+1/4
1+1/4
2+UP
2+1/8
2+1/8
3+UP+UP+STAR
3+1/16
3+1/16
4+UP+UP+UP
Blocked Corridor
Unblocked Corridor
Blocked Corridor leading to a one point gote play

The chilled values for some simple corridors.
Adapted from Wolfe.
Figure 5.1

The topmost tile on the right column illustrates a one point *gote* (pronounced "go'tay") play. As Wolfe defines it [WD1, p. 34]:

> An x point gote is go terminology for the unchilled game { 0 | - x }.

Black's move in this hot game is worth an extra point more than any move by either side at any time in the games shown in the other two columns. However, after Black's move, the game becomes a number, in this case, the number 0. The right column illustrates tiles that lead to a one point gote play. Their value is

> length - 1 + an infinitesimal.

If the length is odd, then the infinitesimal is

> (length - 2) sum of UP's.

If the length is even, then the infinitesimal is

> (length - 2) sum of UP's + *.

To prove these formulas, evaluate the last move and inductively relate length to length + 1.

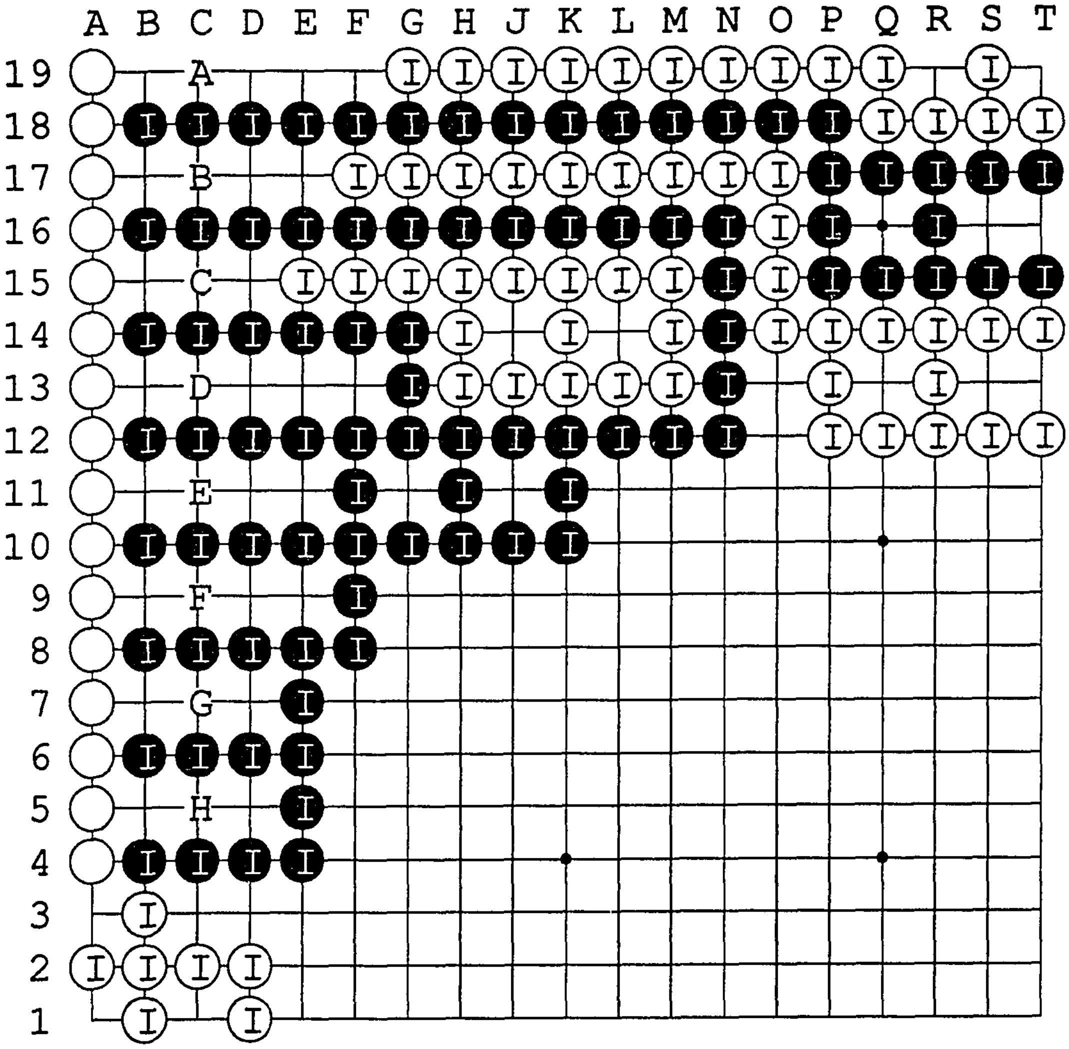
Sockets

In Chapter 3, we defined a chain as the unit of capture, and a group as the unit of life. Chains can be loosely connected or strongly connected. Strongly connected chains can never be prevented from merging into one group (not necessarily one chain) when play within these chains is strictly alternating and there are no kos. Loosely connected chains may be severed, but only at a cost in other parts of the board. Wolfe defines a socket as "A play (eventually) required to connect a group to life" [WD1, p. 68]. In our terminology, a socket is what go players think of as potential cutting points among strongly connected chains where at least one chain is not immortal. It may be helpful to substitute "between two" for "among" in the preceding sentence.

The value of a socket is related to the values of the tiles it will eventually connect with. Figure 5.2 shows Wolfe's first example. The chain of White stones from A19 to A4 has no eyes. The chain of stones in the lower left corner has two eyes and is immortal. For both chains to survive as one group, White must connect at A3. A3 is called the socket for the two chains. Additionally, Wolfe shows that various types of sockets, all uncuttable connections, have different values. A3 has a value of 1, a point for Black since White eventually will make a move here. Wolfe portrays other sockets with ko values in their game tree [WD1, p. 39].

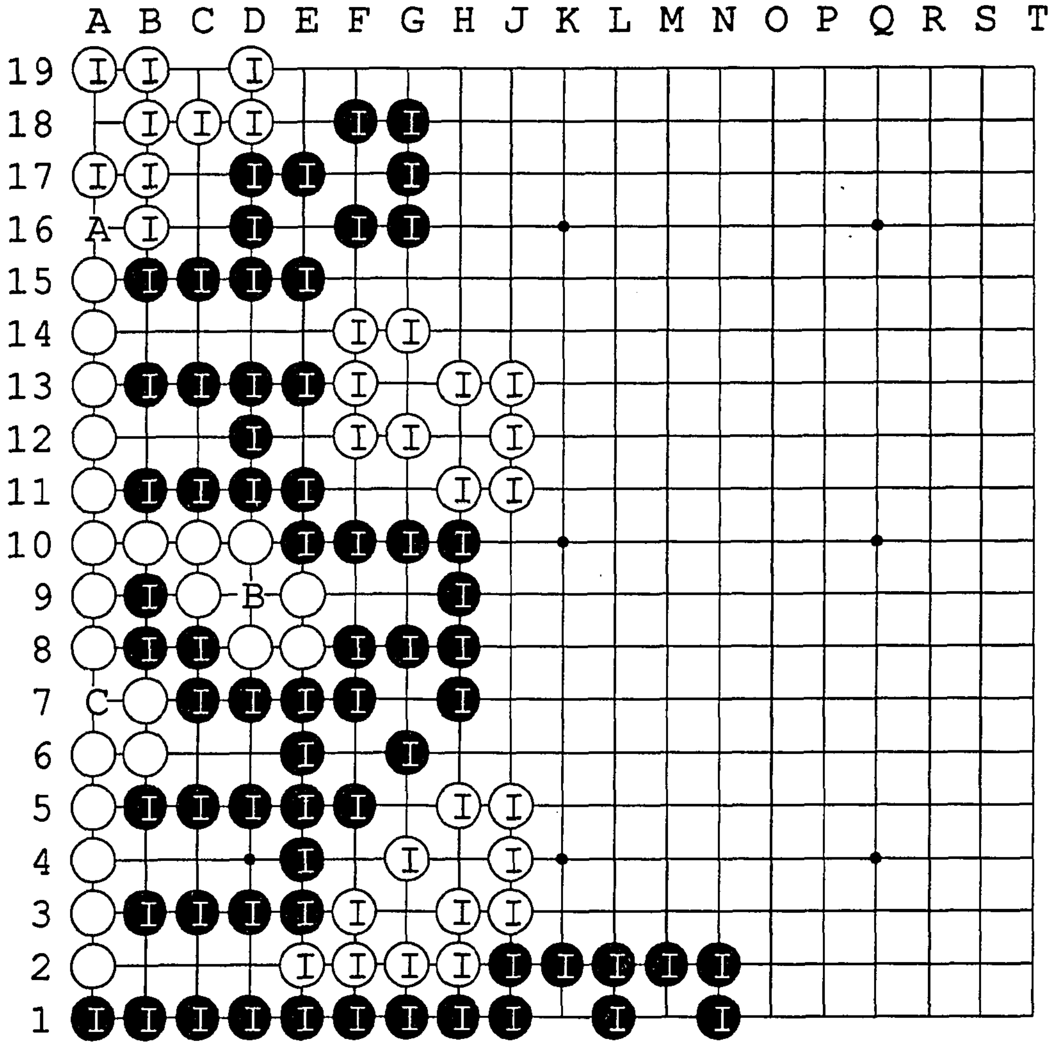
Returning to the rest of Figure 5.2, Tiles A through C are unblocked corridors of length of 5, 4, and 3 respectively. Tiles D through H are blocked corridors of length 5, 4, 4, 3, and 3, respectively. Applying our previously derived formulas, the values for these tiles are: A = 1+1/4, B = 1/2, C = 0, D = 3+1/16, E = 2+1/8, F = 1+1/4, and G = 1+1/4. Remembering that the socket at A3 has value 1, the naive expectation is that the entire value for the tiles on the left (we ignore the immortal groups on the right as notational convenience) is the sum of these 8 terms, namely 12 + 9/16. Wolfe has stated and proven a general theorem [WD1, pp. 36-41] for corridors connected by one socket which he applies here to get the true value for these 7 tiles and 1 socket, namely, 12 + 21/32 + *.

Wolfe generalizes his theorem to multiple sockets and provides an algorithm for solving these positions. Figure 5.3, illustrates a position with three White sockets at A, B, and C that Wolfe solved in his thesis [WD1, p. 47].


A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
A
B
C
D
E
F
G
H
I

Unconnected group invading many corridors, adapted from Wolfe "Mathematics of Go: Chilling Corridors" [WD1, p.38].

Figure 5.2


A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
A
B
C
I

Multiple sockets, adapted from Wolfe [WD1, p. 47]

Figure 5.3

In this chapter, we have seen that Wolfe, has proven Berlekamp's BIG THEOREM, classified types of corridors, proven the value of single socket positions, and given an algorithm for multiple socket positions. Additionally, among other things, Wolfe gives several of Berlekamp's problems, along with their complete solutions, including a "9-dan stumper."

Wolfe has extended Berlekamp's work in several directions. Until now, the author of this thesis has endeavored to clarify the work of Berlekamp and Wolfe. We now aim to apply some of their methods to go and computer go, to bridge the three cultures. The next chapter aims at a precise notion of corresponding outcomes between go and Mathematical Go for simple tiles.

CHAPTER 6. Mathematical Go values and go scoring

When playing Mathematical Go, the sum of the chilled values indicates who gets to play the last move, a Mathematical Go win. Not so obvious is the corresponding outcome in the real game of go. This chapter will examine the relationship between chilled values and the outcome of a real go game. *We shall ignore all games where seki or draw by repetition are part of the best play tree of any tile, unless stated otherwise.*

We shall use the acronyms MGO for the Mathematical Go outcome under Mathematical Go rules, JO for the Japanese outcome under Japanese rules, and CO for the Chinese outcome under Chinese rules. *When using CO, we refer to the playing area only; we ignore the count of immortal stones in any tile, assuming it is balanced elsewhere by their corresponding negative value.* First player refers to the player on move, second player refers to the first player's opponent.

As discussed in Chapter 2, there are four outcome classes in a WW Game: positive is a Black win, negative is a White win, 0 is a second player win, and fuzzy is a first player win. The chilled go game is an extended WW game with another possible outcome; a draw with no result, i.e., by endless repetition. No MGO measures the size of any outcome. There are four outcome classes under JO and CO: Black wins by a measured score, White wins by a measured score, jigo (both measured scores are the same), and a draw with no result, i.e. by endless repetition.

There are four classes of WW values that are applicable to chilled go endgames:

I. Integers

II. Fractions, (including an integer component)

III. Infinitesimals (including a numeric component)

IV. Non-infinitesimal hot games

We have already discussed that *the best selection among a choice of tiles is to play in the highest available class*. We shall discuss how to select a tile from a class and how to score the result of that move.

Play in class I

In this case the underlying go game consists of settled territory and, optionally, neutral points, or dame. The integers of chilled go correspond exactly to the Japanese and Chinese way of scoring for the playing area involved. Sometimes it's notationally convenient to consider 0 as an exception to class 1, and treat it explicitly. When there are no dame present and the net score is 0 the outcome is:

MGO: a second player win. JO and CO: a tie.

As we discussed in Chapter 4, there is no strategy involved here, every play has the same value under MGO. Under JO or CO the game has ended, immortal groups do not have their eyes filled under these rule sets. When there are no dame present and the net score is not 0 (>0 for Black or <0 for White) the outcome is:

MGO and JO and CO: a win for the leading player. For JO and MGO, the magnitude of the score equals the value of the game.

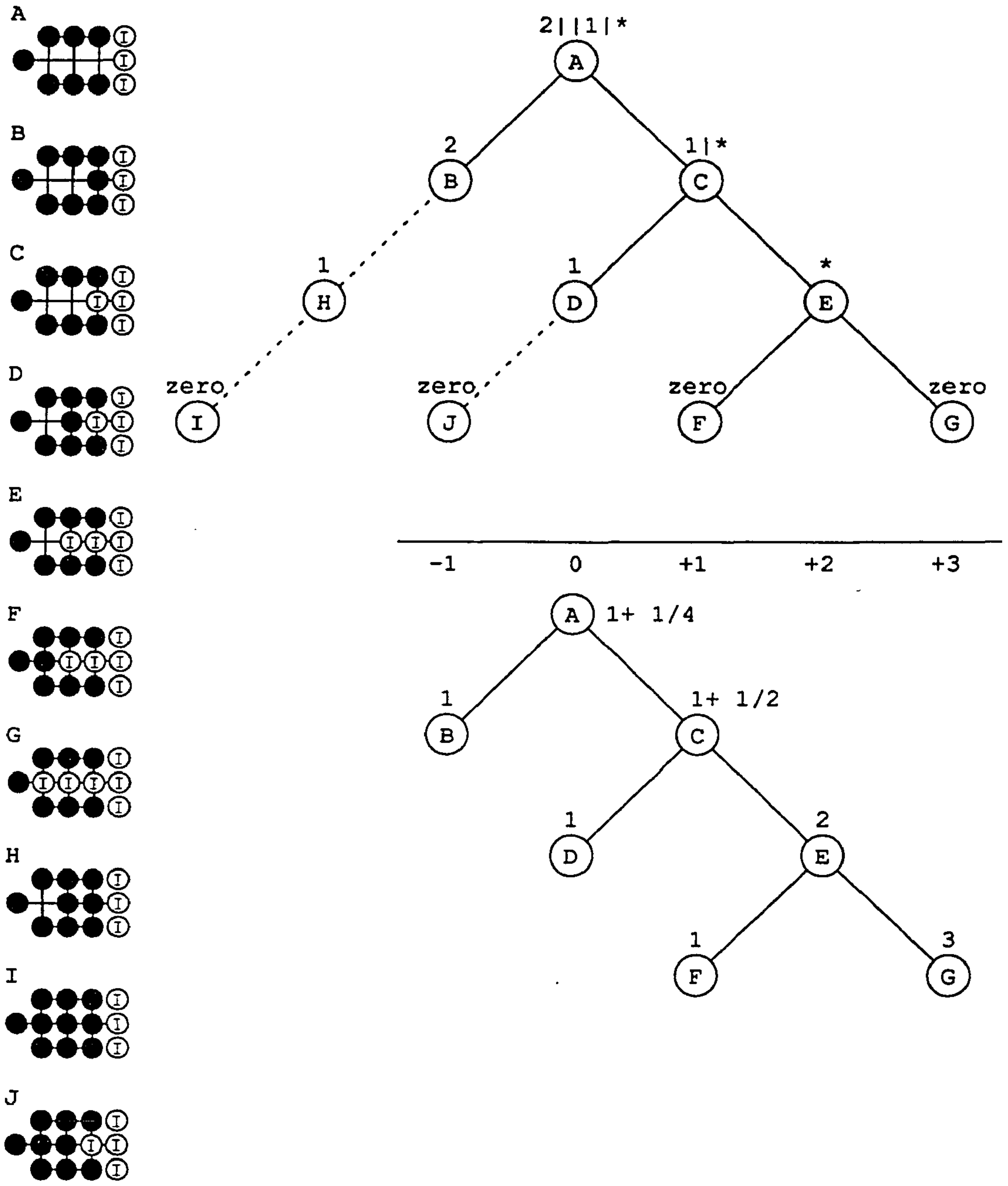
When there are an even number of dame present, the outcomes are the same as the two cases discussed above. Under MGO and CO the players must alternately play on the dame until none are left, else the score will be affected. Under JO the game has effectively ended. When there are an odd number of dame present then MGO and CO treat the dame as 1 extra point with the appropriate sign for the first player. JO ignores the dame, the game has effectively ended. The outcome is determined by the aforementioned rules.

Recall that in Chapter 2, we saw that $* + * = 0$, hence $N \times (* + *) = 0$. The chilled value of a move on dame is $\{ -1 \mid 1 \} = \{ \mid \} = 0$. The incentive for either player to play in dame under MGO is -1. Note that the chilled value of a move on dame is a reversible move, hence won't show in a canonical form tree, but will show in a best play tree. It is moot whether to include the tile 0 in class I or treat it as an exception.

In a sum of tiles containing some tiles other than class I, and, optionally, some class I tiles, best play will be outside class I tiles. Play outside a class I tile will have the same incentive under JO and CO. Play in higher classes of tiles eventually (but not necessarily immediately) leads to a class I tile or 0. No move can result in a change to a higher class.

Play in classes II and I

This situation illustrates a part of Berlekamp's BIG THEOREM that we discussed in Chapter 5. The score of a game after chilling and warming is bounded by 0 or * under JO or MGO but not under CO. Figure 6.1 shows A, a corridor of length 3, A's best play game tree, A's chilled game tree, and its cumulative chilling scale. In the chilled game tree, A and C are


A
B
C
D
E
F
G
H
I
J
2||1|*
A
2
B
1|*
C
1
H
1
D
*
E
zero
I
zero
J
zero
F
zero
G
-1
0
+1
+2
+3
A
1+ 1/4
1
B
1+ 1/2
C
1
D
2
E
1
F
3
G

A chilled corridor of length 3. The best play game tree is on top, the chilled morphism is below. Mappings are not used for play past what originally was a number.

Notice the negative incentives at each move for both players.

Figure 6.1

Class II tiles, the other tiles of the chilled game tree are Class I. Notice that the transition from class II to class I by any path is within the bound previously mentioned. Play within class I tiles form a finite descending chain bounded by 0.

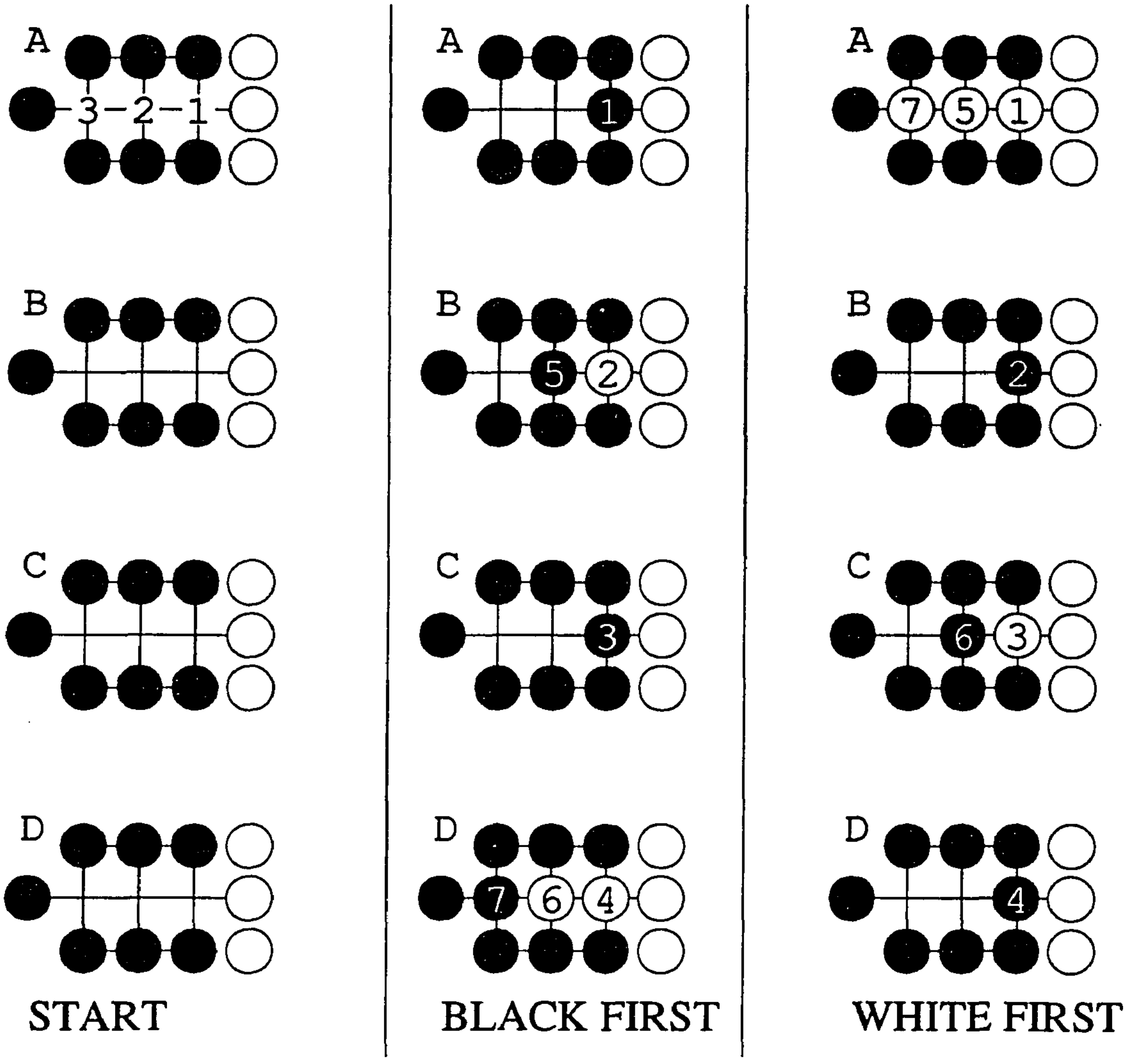
In Chapter 4, we evaluated an individual tile without accounting for moves outside the tile. Similarly, this figure shows the horizontal scale indicating the sign and magnitude of the cumulative chill relative to A. For example, Tile B, as a descendent of tile A, evaluates to 1. Independently, B evaluates to 2. Notice that with strict alternation of moves, only tiles B, C, and D are reachable from A.

If we represent a class II tile as I + F, where I is an integer and F is a proper fraction, then best play within that tile eventually results in class I tile(s) with value I or I + 1 depending on play. Regardless of the magnitude of F, the first player will get the disputed point. Under JO, this means an extra point of territory, under MGO, after chilling and warming, this means an additional move when settling settled territory. See below for examples when F = $\frac{1}{4}$ and F = $\frac{3}{4}$. CO follows this rule, and additionally awards a parity point for the path created during play to that player who has made an extra move in that tile.

If we represent a class II tile G as I + $\frac{N}{D}$, where I is an integer and $\frac{N}{D}$ is a proper fraction, then the value of the sum of D copies of G is D x I + N. If If we balance G with settled White territory -G; then under MGO, whoever plays first wins, while under JO the result is a tie on score. Under CO, in addition to the value of D x I +N, the outcome depends on whether or not one side has played an extra move.

Figure 6.2 illustrates this concept for N = 4, where N is the number of copies of a tile. A, B, C, and D are notational conveniences for copies of the tile we called A in Figure 6.1. In both the middle and right columns, we see alternating play within the sum of g x 4 without any ordering on the tile selection, other than the dictates of good strategy. Under JO, and under MGO the value of A is $1 + \frac{1}{4}$, no matter who moves first. The value of G is 4 x $(1 + \frac{1}{4}) = 5$. Under CO, the value of A is 3 when Black moves first or 1 when White moves first, taken in isolation from a larger game. The outcome of G, i.e. 4 x A under CO differs from from MGO and JO, depending on who gets the first move. The AGA [AGA] has resolved this discrepancy between the two methods of go scoring by absolutely barring the first player from making an excess number of moves. To review what we have covered -- multiplication of a fraction by a multiple of its denominator yields an integer. When multiplication does not yield an integer, the size of the fraction will not affect the outcome. CO scoring includes the parity of moves made on the playing field before determining the outcome.

Using the numeric markings for the playing field of Tile A in the left column, after selecting a tile, proper play is 1 before 2 before 3. An example of a poor strategy by Black, White playing first, would be to counter every invasion at 1 with a blocking move at 2. This poor strategy yields a lower outcome for Black under all three methods of scoring.


A
3
2
1
B
C
D
START
A
1
B
5
2
C
3
D
7
6
4
BLACK FIRST
A
7
5
1
B
2
C
6
3
D
4
WHITE FIRST

Let g, alias A or B or C or D, be a straight corridor of length 3.
The left column shows an initial game G = g+g+g+g.
The middle column shows G played best after Black moves first.
The right column shows G played best after White moves first.
Figure 6.2

If we play N copies of Tile A, the size of Black's win for each scoring system depends on who moves first:

N = 1, Black plays first:

MGO : 2 moves in reserve.

JO : 2 points of territory.

CO : 3 points of area, Black has made an extra move.

N = 1, White plays first:

MGO : 1 move in reserve.

JO : 1 point of territory.

CO : 1 point of area, even number of moves made.

N = 3, Black plays first:

MGO : 4 moves in reserve, Black fills the last dame.

JO : 4 points of territory, dame can be left unfilled.

CO : 5 points of area, Black has made an extra move.

N = 3, White plays first:

MGO : 3 moves in reserve, Black fills the last dame.

JO : 3 points of territory.

CO : 3 points of area, even number of moves made.

N = 4, Black plays first:

MGO : 5 moves in reserve, Black fills the last dame.

JO : 5 points of territory, dame can be left unfilled.

CO : 6 points of area, Black has made an extra move.

N = 4, White plays first:

MGO : 5 move in reserve, Black fills the last dame.

JO : 5 points of territory, dame can be left unfilled.

CO : 4 points of area, White has made an extra move.

N = 8, Black plays first:

MGO : 10 moves in reserve. players alternate filling dame.

JO : 10 points of territory, dame can be left unfilled.

CO : 10 points of area, even number of moves made.

N = 8, White plays first:

MGO : 10 moves in reserve. players alternate filling dame.

JO : 10 points of territory, dame can be left unfilled.

CO : 10 points of area, even number of moves made.

N = 8 million, Black plays first:

MGO : 10 million moves in reserve. players alternate filling dame.

JO : 10 million points of territory, dame can be left unfilled.

CO : 10 million points of area, even number of moves made.

N = 8 million, White plays first:

MGO : 10 million moves in reserve. players alternate filling dame.

JO : 10 million points of territory, dame can be left unfilled.

CO : 10 million points of area, even number of moves made.

In all 3 forms of scoring, the results are identical when playing in 2*N games when N is the denominator of the value of the game.

Play in classes III, II, and I

To determine the outcome of a sum of class III tiles, we first review some terminology. A stop is "any game which is a number" [WD1, p. 68]. A stopping position from a tile whose value is a hot game *is the first tile in the path of a strictly alternating sequence of moves which is a stop.* A stopping value is a pair, the value of the stop and the player on move after the stop. Each tile which is a hot game has a Black stop and a White stop, according to who moves first from the tile. This is similar to comparing JO and CO for class II tiles.

Refer back to the chilled game shown in Figure 4.6. B is a hot game. The Black stop from B is (B1, White), the value of B1 with White to move. The White stop from B is (B3, White), the value of B3 with White to move. In class III games, the value portion of the Black and White stopping positions are always equal. This follows directly from the definition of infinitesimal. The average of the value of these two identical stops, which is another way of saying the value of either the Black or White stops, is called the mean value of B.

Let us assume the sum has been calculated. If the sum is 0 then the MGO is a second player win and the JO or CO is a tie. If the sum is positive, Black has an MGO win, under JO or CO; Black to play has a 1 point win, White to play has a tie. If the sum is negative, White always has an MGO win; under JO or CO, Black to play has a tie, White to play has a one point win. If the sum is fuzzy, the outcome depends on the stopping value, When the stopping value is an integer, under MGO the first player wins; under JO or CO, the first player picks up a point. If the stopping value is a fraction, complexity arises. In Chapter 4, we formulated step four as "play out the game under WW techniques." We mentioned that Wolfe has listed the lattice for all tiles occurring in corridors. To elaborate, the technique here is a two stage process, First select the tile with the greatest incentive, then make the play that determined that incentive. In playing out the game, it is not necessary to know the final outcome before making a move. *Indeed, we are not aware of any published algorithm for quickly determining either the partial ordering or the sum of infinitesimals for all occurrences in Mathematical Go tiles.*

We speculate that the cost of making a large number of constructive comparisons needed to determine the partial ordering and sum may be near the complexity of creating a non-WW exhaustive search tree. Chapter 7 will discuss the data-dictionary approach. based on dynamic programming principles.

An infinitesimal is closer to zero than any number. In Figure 5.1, on the rightmost column, the blocked corridor of length 3 leading to a one point gote play has a value of 2 + UP. If we played one large game consisting of one million copies of this tile under JO or CO, Black playing first would win by two million and one; White playing first would lose by two million. *This is unlike multiplication for class II tiles which obey the laws of normal arithmetic under MGO and JO.* This type of computation is similar to the go strategy of miai which we discuss in Chapter 7.

A fundamental metric in a Mathematical Go position is the excess of UPs or DOWNs summed over all the tiles. As a first approximation, the proper strategy for Black is to remove the largest number of DOWNs and only when no DOWNS are present, to remove the fewest number of UPs. The dual strategy applies for White. As a first approximation, the net number of UPs and DOWNs predict the outcome. An excess of UPs is positive, an excess of DOWNs is negative. When the number of UPS equals the number of DOWNS, lower-order infinitesimals (e.g., tiny and miny) will effect the outcome.

Play in classes IV, III, II and I

When class IV values are involved, the analysis is more complicated. Games with class IV values, that are numbers after one move is made by either player, are known as switch games. Let G = { B | W }, where both B and W are numbers and B > W by a finite amount. As discussed in Chapter 4, the value $\frac{B+W}{2}$ is called the "mean value" of G, $\frac{B-W}{2}$ is the temperature of G, and the actual value of G is fuzzy between B and W. There is a simple algorithm for best play in a sum of switch games. The algorithm is "*be greedy,*" i.e., always play in the switch with the greatest temperature. With N switch games, N moves are needed to remove all Class IV values. Then, follow the strategies discussed above.

In the general case, we are not aware of any published algorithm for determining exact outcomes within anything short of the exhaustive game tree. However, WW [BE1, pp. 159-163] gives an algorithm called "thermostrat" which provides for near-optimal play. The name "thermostrat" stems from thermal strategy, i.e., the strategy of playing in hot games. The thermostrat algorithm guarantees its result to be bounded by the maximum temperature of any one tile from the set of all tiles in the large game. Consider a game G with a million games $g_1, g_2, \ldots, g_{1{,}000{,}000}$, such that the maximum temperature for any g_i is 10 and any game may be identical with or different from any other game. WW proves that a player, using thermostrat in G, is guaranteed to make at most 10 suboptimal moves to a final stopping position within 10 of the optimal stopping position, even though the compound game may last many millions of moves.

Without going into unnecessary technicalities, thermostrat uses graphic methods, called "thermographs," to depict games and their numeric values when cooled by different temperatures. The x-axis is the number line, the positive y-axis is increasing temperature. Stopping points form the leaf nodes along the x-axis. In the switch game $\{m \mid n\}$, the thermograph is an isosceles right triangle with hypotenuse from m to n and height $t(\{m \mid n\})$. WW gives rules for calculating compound thermographs. In a bottom-up fashion the compound thermograph is calculated to discover the "ambient temperature," a property of the width of the compound thermograph. Thermostrat dictates that a player should select the component that is widest at the ambient temperature; then the player should move in the option of that component, that determined the player's boundary of the thermograph of the selected component at that temperature. *According to thermostrat, and unlike the aforementioned greedy algorithm, a player moves in the ambient temperature, which often is different from the hottest game.* Thermostrat works extremely well in large games with many components.

This chapter has presented the metrics for making a move in a Mathematical Go problem. Among classes select the highest class to play in. Within a class, select the tile with the greatest incentive. Within a tile, select the indicated move from the best play game tree. Calculating the final outcome is not necessary for selecting the best move. This chapter has compared how positions are scored with MGO, JO, and CO.

In Chapter 7 we discuss how to enumerate distinct tiles small enough for any one to fit on a go board, we discuss how such a list may be used for designing problems, and we discuss how such an enumeration is impractical, given today's technology.

CHAPTER 7. Design issues in Mathematical Go

How might one proceed in composing an interesting Mathematical Go problem? What would make a Mathematical Go problem interesting to a go player?

One esthetic attribute of a problem is that it has instructional value, that it teaches some theme, some useful pattern, shape or tactic. This thing being taught need not be directly related to the game.

Regarding the instructional value of his thesis, Wolfe states [WD1, p.1]:

> Go players can find quicker ways to improve their game than to read this thesis. Many of the positions analyzed do not tend to come up in contexts which take full advantage of the subtleties of the results. However, there are some wonderful lessons for the go player.

Wolfe is a very strong mathematician who does not state his go playing strength. David Mechener, a former professional go student in Japan, now a mathematician whose strength is unspecified states concerning tsume-go (endgame puzzles similar to Mathematical Go) [MD]:

> The central point of tsume-go is not to memorize corner shapes or learn how to kill groups; it is to sharpen the precision of your reading and to discipline your mind.... Confucious said, "It is better to properly read out one problem then to flip through all the tsume-go in Japan" -- well, maybe he didn't but he would have if he'd been a strong player.

Nagahara, a 5-dan professional go player, writing about go problems, states [NY, p. 59]:

> The reader should not be discouraged by failure to solve the problems. The real benefit comes from thinking seriously about each and then comparing one's answer with the solution provided. This is the best way to discover and correct your own blind spots.

Ikeda, a keen amateur go player and one of the main figures in the development of Japan's computer industry states [IT, pp. 14-15]:

> There is little chance that we will ever know when, by whom, or in what way the game of go was created, but at least we can use our imaginations....
>
> The explanation offered by Go Seigen (9 dan) is that the go board and stones were used in the past as tools for research into divination and other fields of learning, or for presenting findings in these fields, but after the invention of paper they gradually evolved into equipment for playing a game. One can imagine diviners or philosophers tiring from their researches with the board and stones and suggesting to their fellows that they play a stone-placing game.

Another esthetic attribute of a problem is that it is difficult. Wolfe [WD1, p.3] displays a Mathematical Go problem that has stumped a 9-dan professional. Wolfe describes the complexity of the problem: "All of the first twenty or more moves are equally difficult, and on each of these moves, few choices lead to a winning position." He adds that many strong amateur go players who were shown this same problem said the problem looks easy. He continues: "even if a go player thinks he has the answer, verifying it requires an opponent well versed in the theory."

In contrast to the 9-dan stumper where every move on both sides requires a difficult selection from many choices, the endgame problems in checkers, chess, or traditional go are typically predicated on discovering one or at most several key moves at the beginning of the solution. Whether these key moves be forcing or waiting, the solution typically follows a certain theme and will not require exhaustive analysis of side branches. The second player's key counter-measures are typically few in number.

Another esthetic attribute of a problem is that it look natural, as if it might arise in play. Most Mathematical Go problems would appear artificial to a go player, but then so do many traditional go endgame problems.

A composed Mathematical Go problem can be tested and/or refined by playing it out. A typical test might proceed as follows. The problem giver (human or computer), knowledgeable in Mathematical Go strategy, prepares the exhaustive solution before the contest begins. The problem solver (a human go player) should be unfamiliar with Mathematical Go strategy. That is, go knowledge should be pitted against combinatoric game theory. The problem solver should first take the purported winning side, and try to demonstrate success against selected defenses by the problem giver. After several failures, the problem giver can reverse sides and show at least one winning variation. Then, the problem solver could resume playing the winning side trying to discover other winning lines. Role reversals would continue to take place until the problem solver learns to win against all counter-measures of the problem giver. It would be up to the gamesmanship of the problem giver to hide the best moves for as many iterations as possible.

In go, the accepted technique for working classical problems is to uncover all the variations to a problem in one's head without touching the stones until the final solution with all its variations has been grasped. This process of thought and concentration is called "reading." It is considered a prerequisite for increasing one's go strength. What Mathematical Go offers to the go player is an alternative map to problem solving. For example, consider a go player learning the solution to Berlekamp's problem discussed in Chapter 4. There are 17 moves, including captures, hence more than 17! possible sequences. The go player might guess most of the best play moves within a tile, but learning the selection criteria using the role reversal method just described would be long and arduous. There might always be a discrepancy within the player's mind between what relevant sequences have been learned and what other relevant sequences must be learned before confidence in mastering the solution is acquired.

The problem giver has a much simpler task. Figure 4.10 is the mathematically proven complete and correct solution to the problem. It is a road map through tortuous terrain. Wolfe [WD1] gives several such maps for the more difficult problems presented in his thesis. Some of these are too hard for a weak player to memorize. If we change the scenario above to allow the human problem giver access to his maps (described in Mathematical Go semantics) and the problem solver access to writing down all the variations encountered, the advantage of Mathematical Go over go knowledge would be even more dramatic. For the computer problem giver, the advantage of Mathematical Go maps over exhaustive solutions is a huge savings in space.

How does one create difficult Mathematical Go problems? Looking at a Mathematical Go solution and the meta-strategy involved should provide some clues. Chapter 6 showed that a normalized positive sum of games will allow Black as first player a one point win. If during any two move exchange, Black misplays and the sum is no longer positive, then Black will get a JO draw, but an MGO win if the sum is now zero. Chapter 6 also discussed how best play selects tiles in stages, from class IV values down to class I values. We saw in the trees of class III tiles that incentives are often asymmetric (see Figure 4.6). Play becomes difficult when there are positive and negative tiles such that the incentives can be matched up. Additionally, play is more difficult when topologically distinct tiles have identical or opposite canonical values. For example, see the left and middle columns in Figure 5.1. In general, a difficult Mathematical Go problem will have the normalized values of the component tiles balance, except for a small infinitesimal; furthermore, play in these infinitesimals should keep the position in balance as the game is played out.

Nagahara, in "Strategic Concepts of Go", lists "miai" as his first strategic concept. He proceeds to explain [NY, pp. 3-4]:

> Miai means "seeing together." It refers to two points which are related in such a way that if one of them is occupied by a player, his opponent can handle the situation by taking the other....
>
> An important point to notice about miai is that the two moves are not often urgent. That is, they are in a state of equilibrium.

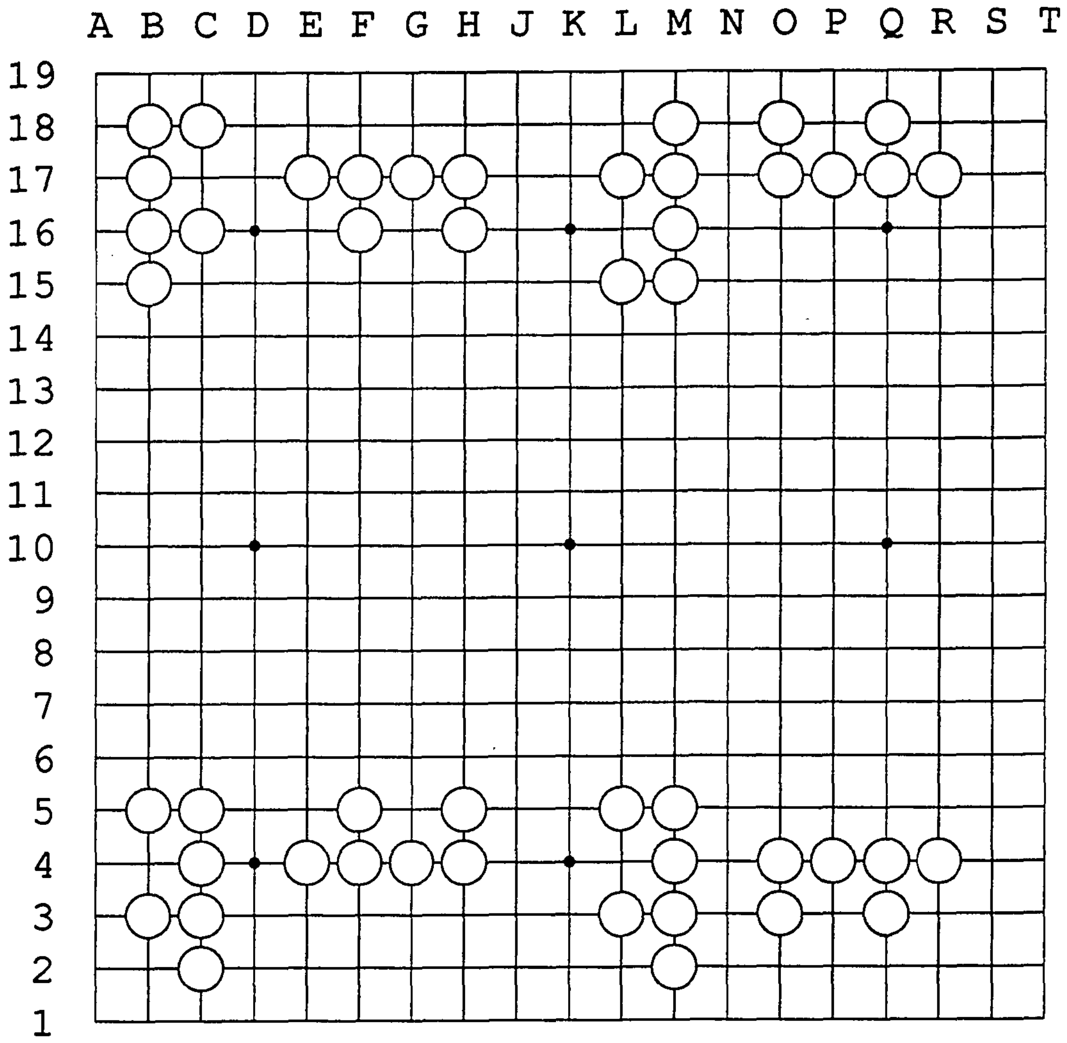
In WW [BE1], the first strategy discussed is the "Tweedledum and Tweedledee Strategy." The strategy applies in a game where every component G_i has a corresponding negative $-G_i$. The second player copies the first player's move, but in the corresponding negative component. Thus, after every pair of moves, the sum of the remaining games stays at zero until all moves are exhausted, the first player has no legal move, and the second player wins.

All published Mathematical Go problems utilize these concepts when selecting the composition of tiles to go into a problem. Tiles and their negatives are chosen. Tiles that after several moves match their negatives are chosen. If the overall sum is to remain positive (or negative for White after every two moves), then there need be many examples of miai that keep that advantage. If there exists one very large incentive then it will be played first. After the first key move, the incentive of the second player's threat is very often the incentive of the first player's reply. As we discussed before, a game often starts hot and chills until everything is settled. The problem solver can reverse roles many times during a problem session, and consistently lose, almost always by one point. *Many small mistakes can be made along the way before another whole point is lost.* Give a 1 and $\frac{1}{2}$ komi to the problem solver, and his performance would benefit enormously. The naive idea that Mathematical Go problems are easy stems from the observation that one small mistake made by the problem solver will result in losing by one point, yet many large mistakes by the problem solver will result in the same loss by one point.

The problem solver's goal is to recognize the value of local tiles, then select the best move. The problem designer's goal is to obfuscate the solution by selecting tiles whose values run counter to a go player's intuition. Also desirable is to select tiles that are morphologically different, yet the same under the chilling morphism. There are two aids for the problem designer that help determine the value of a listed tile. Wolfe [WD2] has written a toolkit in C under UNIX that performs most of the WW computations on games including games with simple kos. Raymond Chen [CR] has ported Wolfe's toolkit to the IBM PC, running under DOS.

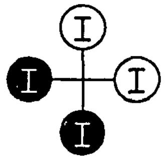
We suggest an algorithm for generating a database of interesting tiles for small playing areas. This is far from an exhaustive method. Two opposing constraints are the desire to be lookup efficient versus the desire to be space efficient. In building a database of tiles and tile values, one might include a graphical user interface to allow input and output via a go board on a monitor. When trying to recognize patterns, the first simplification would be to recognize identical positions save for rotation, reflection, inversion, and color. Figure 7.1 depicts how the same chain of six stones can look like eight different groups to an input grid. There are many approaches to pattern recognition of go board shapes. Mark Boon [BM], who has written the 1991 World Champion Computer Go program, Goliath, describes his pattern recognition algorithm as testing for each pattern eight times to avoid later having to check for symmetry. That is, he stores eight input patterns for each desired shape in his data base. David Fotland, who has written the North American Champion Computer Go Program, The Many Faces of Go, uses multiple representations of the playing board and only one pattern in the database. That is, each move entered at the user interface is interpreted internally as being made on eight boards with the appropriate symmetries. Additionally both authors can flip the bits of their internal board representation to match a position or its color mirror-image against a stored pattern.

We suggest a nomenclature for all tiles, based on their morphology. We preface all tiles by "T" and use the first index to show the size of the playing area. Figure 7.2 shows the T1 tile and the five T2 tiles. We ignore settled territory. We ignore all symmetries discussed in Figure 7.1 and symmetry of color. Any playing area may have potentially nine different values depending on its perimeter. We first surround a playing area in the center of the board, then examine the resulting positions when any one edge or adjacent two edges are removed. This can be likened to sliding a tile from the center to each of the four edges and four corners such that the immortal stones on the perimeter fall off while the playing area remains on the board. Of course, when certain playing areas are placed on an edge(s) it may be impossible to form a legal perimeter. For example, a concave area may border a space too small to form a living group. Chapter 3 illustrates the smallest space needed to form a living group for center, edge,

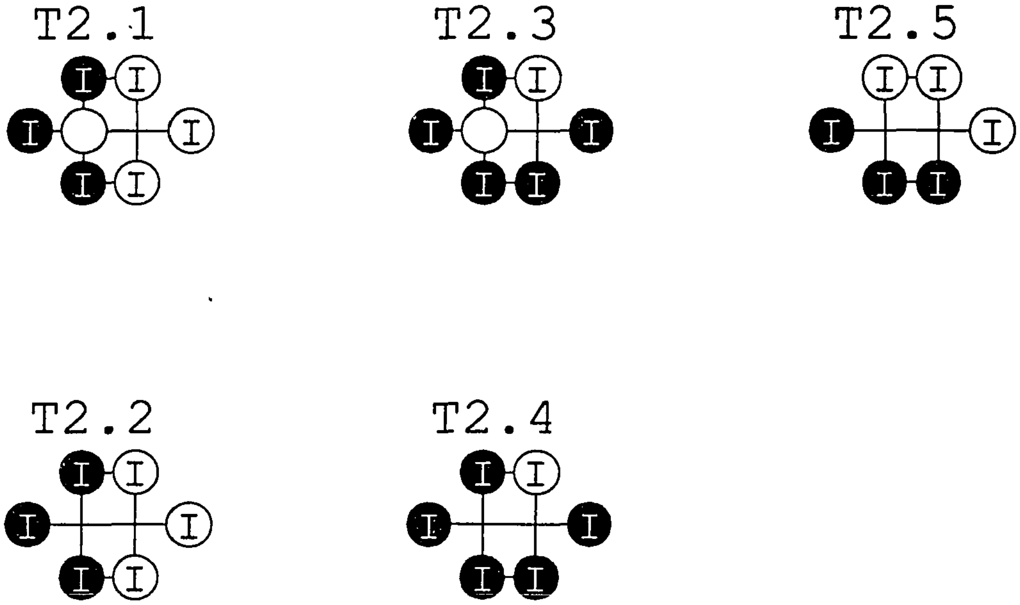

A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

Eight symmetric placements for six stones of one color.
Figure 7.1

T1.1



The T1 tile.


T2.1
T2.3
T2.5
T2.2
T2.4

All the T2 tiles.

All the tiles with playing areas of size 1 or 2;
ignoring the symmetries of Figure 7.1,
non-impacting immortal stone placements,
non-impacting center/edge/corner placements,
non-impacting topological variants,
and settled territory.

Figure 7.2

and corner placements.

Each point in a playing area is adjacent to 2, 3, or 4 other points. Each playing area point can be classified by the different types of immortal stones it contacts:

B - contacts Black, not White, and possibly empty points.

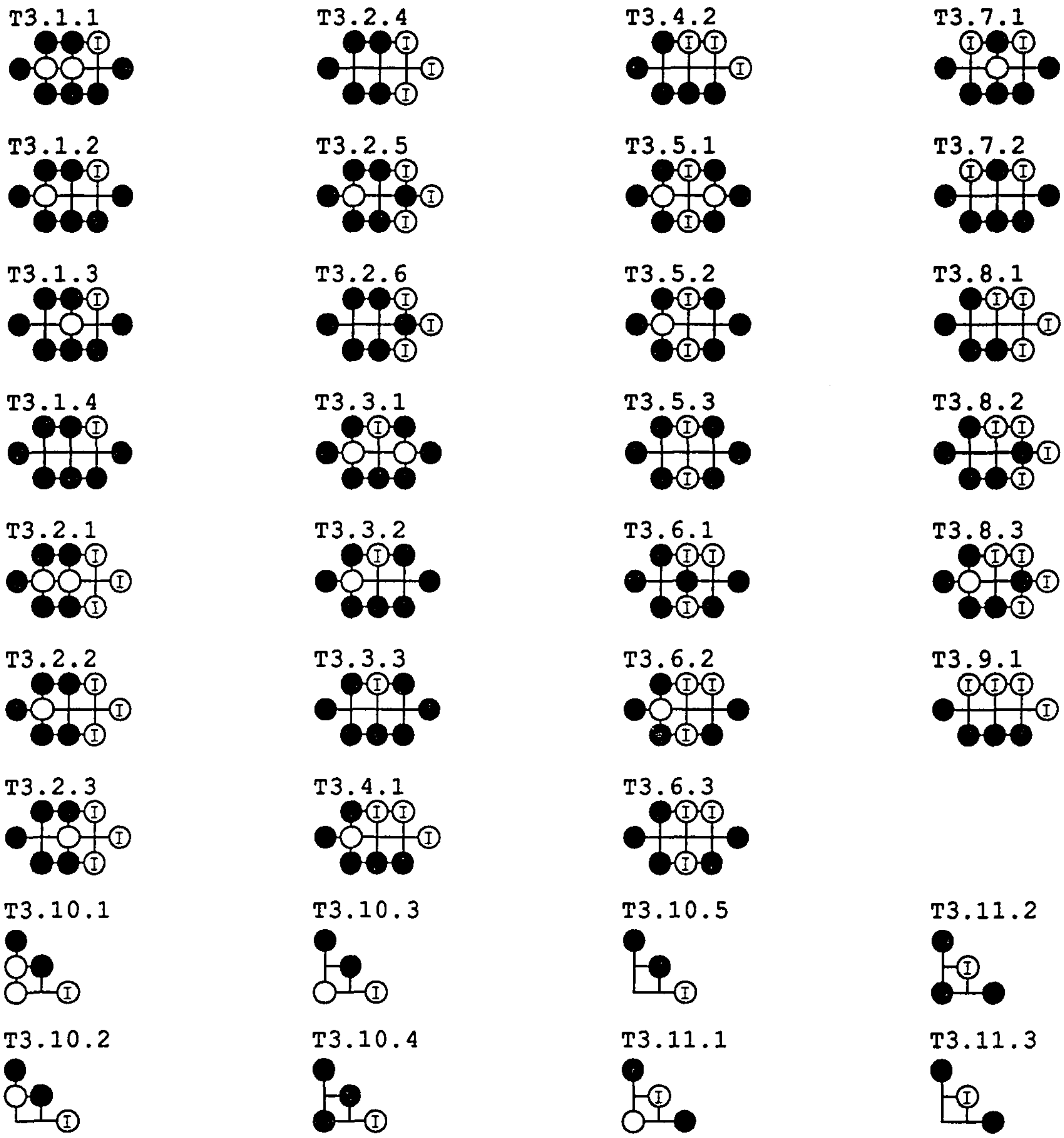
W - contacts White, not Black, and possibly empty points.

S - contacts Black and White, and possibly empty points.

E - contacts only empty points.

To generate all distinct tiles, playing areas are first located centrally. Next, each point is surrounded by the various combinations. Figure 7.3 lists all tiles of playing area 3. We have used a second index to depict playing area relationships. A property of this index is that a tile labeled T3.n.x will have legal moves either to T3.n.y or to T2.z. An ordered string of the playing area point values will determine a unique T3 tile. The only difference among tiles with the same playing area configuration is the presence or absence of stones that are not immortal. For playing areas of size 4 and above it is convenient to add another index based on the underlying topological shapes. Figure 7.4 lists the topological shapes for size 4. For larger sizes, the class E points also need to be partitioned by shape.

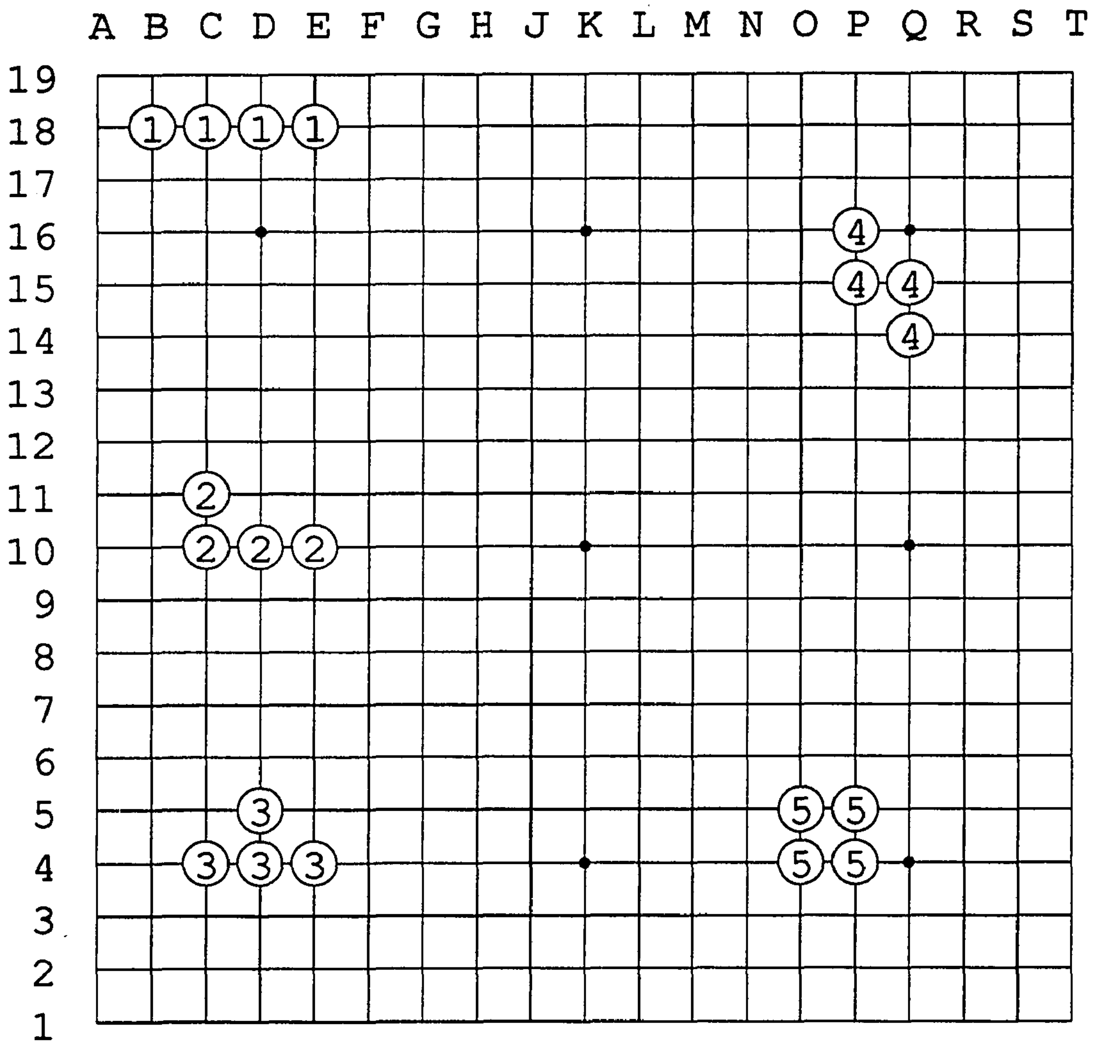
The general algorithm to enumerate all tiles is:


T3.1.1
T3.1.2
T3.1.3
T3.1.4
T3.2.1
T3.2.2
T3.2.3
T3.2.4
T3.2.5
T3.2.6
T3.3.1
T3.3.2
T3.3.3
T3.4.1
T3.4.2
T3.5.1
T3.5.2
T3.5.3
T3.6.1
T3.6.2
T3.6.3
T3.7.1
T3.7.2
T3.8.1
T3.8.2
T3.8.3
T3.9.1
T3.10.1
T3.10.2
T3.10.3
T3.10.4
T3.10.5
T3.11.1
T3.11.2
T3.11.3

All the tiles with playing areas of size 3;
With the same restrictions mentioned in Figure 7.2.

The playing area configuration of n in T3.n.x is:
1 = BBS, 2 = BBW, 3 = BSB, 4 = BSS, 5 = BWB,
6 = BWS, 7 = SBS, 8 = BSW, 9 = SSS, 10 = BES, 11 = SES

Figure 7.3


A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
1
2
3
4
5

The five topological shapes of size 4.
Figure 7.4

```
initialize a database to empty.
for each playing area size (e.g. 1, 2, ...)
    for each internal topology (e.g. see Figure 7.4)
        for each border orientation (central, 4 edges, 4 corners)
            for each playing area configuration (string of B, W, S, or E)
                if the tile has not already been entered in the list
                    of distinct tiles in the database
                    and the tile is legal (immortal stones can
                    be constructed as a living group)
                then
                    add it to the database.
                    add all distinct combinations of potential
                      prisoners in the playing area to the database.
                else if the position is not legal
                then
                    discard it
                else
                    store the tile in a list of symmetrical tiles,
                    store a pointer to its match.
                end if
            end for
        end for
    end for
end for
```

When considering the nine border orientations, there are many instances of duplicate tiles with the same playing area configuration. For example, when tiles T3.1.x through T3.9.x are placed along the left edge, the leftmost immortal black stone is replaced by the left edge of the board. In all these cases, it is obvious that the leftmost playing field point remains in class B, since the two adjacent immortal stones are both Black.

Having available a database of values, the problem designer follows the esthetic considerations discussed above, make the problem instructional, difficult, and natural looking. A good problem should make the underlying Mathematical Go mapping run counter to the intuition of the go playing problem solver. Tiles that are corridors are easy to compose problems with because there is a lot of surface area to share with other tiles. In Wolfe's version of Figure 5.2 [WD1], the stones to the right of column G were not in the diagram, immortality was designated by a line going off the edge of a partial go board. We connected all the immortal stones to live groups. The result is unsatisfying to a classical go player because it is too unnatural. As the size of tiles increases, it may not be possible to create immortal groups that fit within the confines of the 19 x 19 board.

Mathematical Go problems look artificial because they are artificial. From the composer's perspective, corridors are easy to pack into the confines of the board and meet the criteria of difficulty. However, from a go player's perspective, corridors are two walls of same colored stones separated by a width of one. The Japanese term "tewari" is used for retrograde analysis of play. A stone or several stones are selectively removed from a position. From the abbreviated position one imagines that the removed stones are now replaced as if they were the last moves played. If these moves are not efficient then the whole line of play is questioned. A fundamental rule of go strategy is to avoid playing near a wall. A corridor may arise naturally in play when squeezed between two larger areas. However, it is a shape that is not desirable, since the object of the game is to surround large areas efficiently. An individual wall radiates power. Walls facing each other separated by one line of empty spaces will not arise in normal play. Yet to compose a problem with very few tiles interspersed among normal go patterns will significantly detract from the complexity of a problem.

A strong point of Mathematical Go is that it greatly diminishes the amount of computation needed to analyze precisely the results of human play to the very last point. Berlekamp [BE3] has checked the very late endgames between strong professionals to see if they followed Mathematical Go techniques. *He has not found an example where Mathematical Go detected a better play than the actual game.* There is no precise point when a go game definitively becomes a Mathematical Go game. As long as there are issues of life and death or seki or kos greater than one point, Mathematical Go does not apply. In general, for ko-free positions, Mathematical Go is claimed to be absolutely accurate up to games of temperature one; that is, each move will affect the score under JO by less than two points. Wolfe has analyzed completely some positions with temperature greater than one; he has calculated values for rooms [WD1, pp. 26-29], what we have called tiles with one S orientation in the midst of all B orientations.

Go World [GW] provides a collection of records of strong professional games. In reviewing many of these games, we observed that the very late endgame often includes large amorphous neutral shapes that don't neatly fall into any category. When the board does become all tiles, the values of the tiles are quite simple. The benchmark for the superiority of Mathematical Go would be to find an instance where one side could have had a win with Mathematical Go analysis but subsequently lost the game. Professional games have a komi (see Chapter 3) of some integer points plus a half point. Hence, no tie scores are possible. Those games that are settled by a half point can be considered the practical test of the player's awareness of go sense equal to Mathematical Go because only in these contests will they necessarily be playing correctly to the last move. Professional go players frequently estimate the score during play. If a player believes he is winning by some measure, he may steer the game into simpler channels, knowing by his reading that he absolutely will win, even if the measure of the win is smaller. If player A believes he is losing by some measure, and if he doesn't feel like resigning, A is honor-bound to complicate the position as much as he can, hoping that his opponent B will go astray in the murky variations before A will. People who play speed chess call this tactic "always hold out for one last swindle." Thus, if a player saw his lead drop from 5 + 1/2 to 1/2 while not following Mathematical Go, he may have read out the position enough to believe he *knows* one method of winning, why face complications. When a player believes he is 1/2 point behind struggles to get a win and winds up losing by much more, his approach is proper. He may have read out the Mathematical Go solution and rejected a small loss for the risk of winning. In other words, a player's goal is to win. If his reading of the Mathematical Go solution yields an unfavorable result, he may deliberately select a different strategy.

In this chapter we have stated some criteria for designing Mathematical Go problems, we have discussed a design for a program to aid in problem composition, and we have shown why the subtleties of Mathematical Go may not occur in normal play.

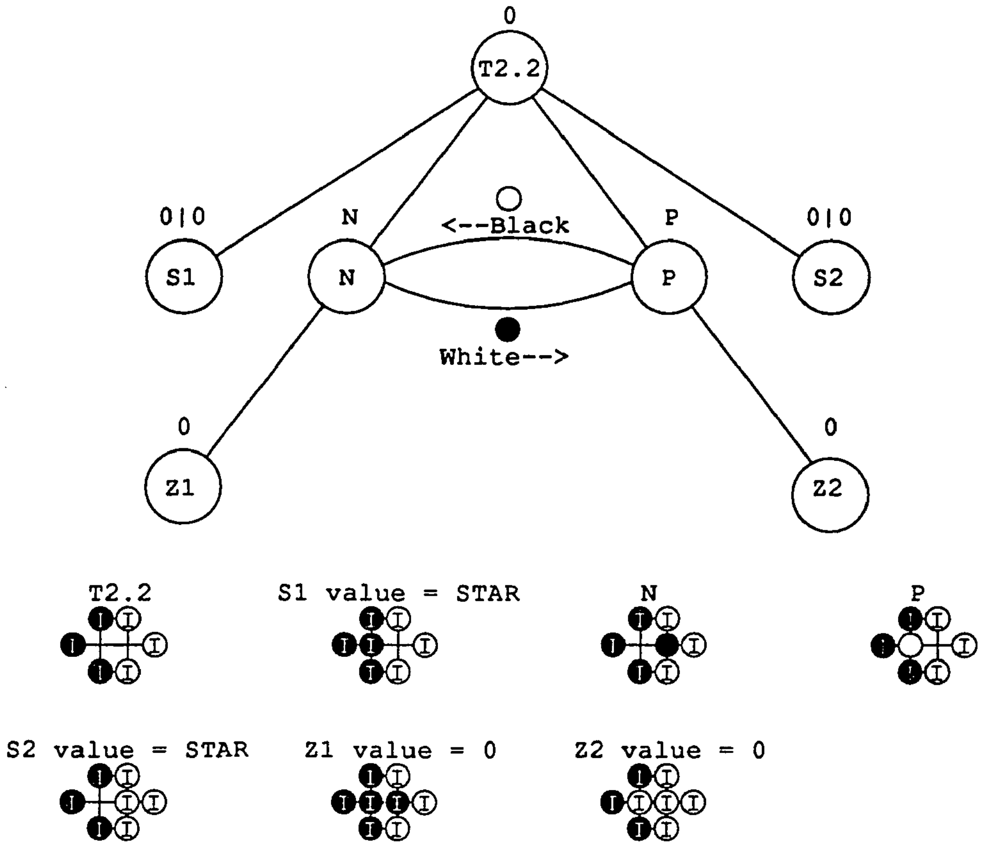
CHAPTER 8. Ko

One fundamental difference between Mathematical Go and go is that Mathematical Go allows immediate recapture of a ko position. This chapter will look at ko, its value, its outcome, and its selection as an option under both Mathematical Go and go rules. Figure 8.1 shows the nearly complete game tree for what we labeled tile T2.2 in Chapter 7. The first player to select this tile has the choice to play in star or in ko. We choose to label the tile P for the mnemonic "P"ositive ko, and the tile N for the mnemonic "N"egative ko. Berlekamp [BE2, p. 78] defines the complement of a game as its color mirror-image. In the sum of a game and its complement, the second player will always have either a win or a draw. When the sum of a game and its complement always lead to a second player win, a value of 0, then each is said to be the negative of the other as defined earlier in Chapter 2. P and N are each other's complement but not each other's negative.

According to Wolfe [WD1, p.60]:

> Berlekamp bounds the value [of P] by: $* \le P \le \{ *+1 \mid 0 \}$.
>
> The lower bound is obtained by assuming [Black] can never win the ko, and the upper bound is when [Black] can always win the ko. So in the chilled game, $0 \le P \le \frac{1}{2}$. Hence, one point kos will not be played until the end of the game, and their value is limited.

Note that T2.2 represents what go players call a "one point" ko -- a ko on which no larger consequences hinge. We will, for the most part, restrict our analysis to one-point kos, as *larger* kos are normally beyond the analytic power of Mathematical Go.

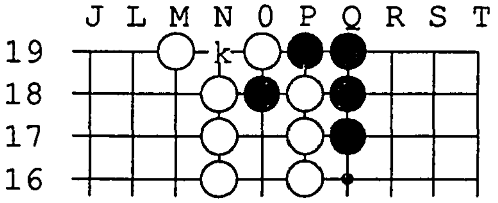

0
T2.2
0|0
S1
N
N
<--Black
P
P
0|0
S2
White-->
0
Z1
0
Z2
T2.2
S1 value = STAR
N
P
S2 value = STAR
Z1 value = 0
Z2 value = 0

The nearly complete game tree for T2.2.

Values shown are WW, not chilled.

(Note: in Berlekamp's notation: P = Ⓚ and N = Ⓚ̄)

Figure 8.1


J L M N O P Q R S T
19
18
17
16
k

A game fragment - Nagy vs. Matsuda

Harry Gonshor Memorial Tournament, 1991.

Figure 8.2

Berlekamp [BE2, p. 94] mentions that P freezes (cools to a number) at temperature $\frac{1}{3}$ to the value $\frac{1}{3}$ + a loopy infinitesimal. We derive this result.

$P = \{ 1 + N \mid 0 \}$

$N = \{ 0 \mid -1 + P \}$

$P = \{ \{ 1 \mid P \} \mid 0 \}$ substituting for N

Let t be the temperature where P becomes a number. Then

$P_t = \{ -t + \{ 1 \mid P \}_t \mid t \}$ definition of cooling

$= \{ -t + \{ -t + 1 \mid t + P_t \} \mid t \}$ cool inner expression

$= \{ \{ -2t + 1 \mid P_t \} \mid t \}$ rearrange numeric terms

$= t + \{ \{ -3t + 1 \mid -t + P_t \} \mid 0 \}$ rearrange numeric terms

When $t = \frac{1}{3}$

$P_t = \frac{1}{3} + \{ \{ 0 \mid -\frac{1}{3} + P_t \} \mid 0 \}$

Let $Q = -\frac{1}{3} + P_t$. Then

$P_t = \frac{1}{3} + \{ \{ 0 \mid Q \} \mid 0 \}$

$= \frac{1}{3} + \{ \{ 0 \mid \{ \{ 0 \mid Q \} \mid 0 \} \mid 0 \}$

$= \frac{1}{3} + \{ \{ 0 \mid \{ \{ 0 \mid \{ \{ 0 \mid Q \} \mid 0 \} \} \mid 0 \} \mid 0 \}$

$= \frac{1}{3}$ + an infinitesimal

For temperatures from $\frac{1}{3}$ to 1 continued cooling only removes the infinitesimal, so chilled $P = \frac{1}{3}$. Using Japanese rules of play and JO scoring, the chilled value is an accurate estimate for predicting the outcome in a multiple of tiles. Figure 8.3 indicates that with either player moving first in the sum of three instances of P, Black wins exactly one point. As we discussed in Chapter 6, Mathematical Go and Chinese scoring differ in the counting of dame.

Outcome of P+P+P under go rules,

Japanese or Chinese scoring.

Strategy A: Black prefers to play in P before N.
Strategy B: Black prefers to play in N before P.

Columns indicate the value of each tile.
Rows show games after the move indicated.
Move 0 is the initial position.
A trailing k means a ko has been taken this move.

Black plays first Japanese scoring

	STRATEGY A			STRATEGY B		
	1	2	3	1	2	3
0:	P	P	P	P	P	P
1:	1+Nk	P	P	1+Nk	P	P
2:	1+N	0	P	1+N	0	P
3:	1+N	0	1+Nk	1	0	P
4:	Pk	0	1+N	1	0	0
5:	P	0	1			
6:	0	0	1			

Black plays first Chinese scoring

	1	2	3	1	2	3
0:	P	P	P	P	P	P
1:	1+Nk	P	P	1+Nk	P	P
2:	1+N	-2	P	1+N	-2	P
3:	1+N	-2	1+Nk	3	-2	P
4:	Pk	-2	1+N	3	-2	-2
5:	P	-2	3	3	-2	-2
6:	-2	-2	3	3	-2	-2

White plays first Japanese scoring

	1	2	3	
0:	P	P	P	Same Result
1:	0	P	P	
2:	0	1+Nk	P	
3:	0	1+N	0	
4:	0	1	0	

White plays first Chinese scoring

	1	2	3	
0:	P	P	P	Same Result
1:	-2	P	P	
2:	-2	1+Nk	P	
3:	-2	1+N	-2	
3:	-2	3	-2	

Figure 8.3

Nevertheless, the technique of playing multiple instances of the same game yields a consistent result with either player playing first, namely P is worth $\frac{1}{3}$

Playing under Mathematical Go rules, ko positions frequently become draws. Let's tabulate the results of some simple games involving P and an integer. Both player's strategy will be to win rather than draw, and to draw rather than lose. We label the outcome win when a player gets the last move, regardless of any territory considerations, or draw if the game repeats eternally.

Game	Outcome	
	Black plays first	White plays first
P + 2	Black win	Black win
P + 1	Draw	Black win
P	Draw	White win
P - 1	Draw	White win
P - 2	Draw	White win
P - 3	White win	White win

From this analysis, it does not appear that the stated range of Berlekamp's P ($0 \leq P \leq \frac{1}{2}$) conforms to the way we interpreted the sum of games in Chapter 6. Predicting the final outcome when kos are present is a complicated matter.

When ko figures in an option for best play, the rules for determining what is dominated or reversible get quite complicated. Figure 8.1 shows that play in tile T2.2 can lead to N or P. Berlekamp's paper does not explicitly discuss this tile, but his presentation of tiles like Tile F of Chapter 4 indicate that he considers the value of this tile 0. Wolfe's toolkit [WD2] computes the value of { *, N | *, P} as 0. A fundamental theorem in Mathematical Go is that G + 0 = G. Black playing first in the Game T2.2 draws by moving to N. More surprising, Black playing first in the game T2.2 -1 draws by moving to N. By adding 0 (T2.2) to a White win, the new outcome is a draw. Once again note that predicting the final outcome when kos are potential options is a complicated matter.

In Chapters 4 through 6, we saw many examples of how Mathematical Go knowledge could be used under go rules to help assess the score of complicated positions, even when certain loopy positions were present. Tile AB of Figure 4.8 leads to a ko after Black moves on point 4 and White plays elsewhere; if White moves on point 4, Black can capture the stones on 3 and 4 by playing on 2. White can reply on 3, capturing the stone at 2 and restoring the position to tile 2. However Black's initial move to 4 is reversible by a White move on point 2, and White's initial move to 4 is dominated by a White initial move to 2. Hence both loopy positions were justifiably eliminated from the best play game tree.

Fundamental to the analysis of a Mathematical Go position is that a tile have a canonical form that is unaffected by the play in other tiles. (This can be qualified by treating tiles connected by sockets as one big tile.) Given a problem setting where tiles subject to simplification exist alongside tiles having potential ko fights; the Mathematical Go simplification of a tile to its canonical form will sometimes be incorrect under JO and CO rule sets. In a ko fight, each player will seek to make ko threats outside the immediate battleground, hoping that his opponent must answer. Recall Figure 2.2. When G ≥ W3, then the move from G to B2 *always* resulted in moves to W3 followed by moves to B5 or B6. The subtrees under B3, B4, W5 and W6 are lost forever. In a ko sequence, let Black make a ko threat from G to B2 with G ≥

W3. If the value of the ko battle is large enough, White will fill the ko and Black can respond in B3 or B4. Additionally, if White reverses B2 to W3, Black may recapture the ko instead of moving directly to B5 or B6. Now White may move to W5 or W6. The nodes B3, B4, W5 and W6, which had been computed to be unreachable are now reachable as responses to ko threats.

Under go rules, positions with kos can take on all the values near the Mathematical Go values we have been discussing. Figure 8.2 illustrates this point. This is a fragment of a game where the ko fight over the points N19 and O19 lasted more than 20 moves. Looking at the points M19-P19, N18, and O18, we have the familiar one-point ko. If Black plays at N19, White plays elsewhere, and Black fills at O19, he will be invading a corridor starting at O17, exactly like those discussed previously. His incentive to win is very near the incentive of winning a one point ko plus the value of a corridor. However, kos can become much more complicated. There are many examples of kos where one side has to invest any number of moves to win a ko. There are also many examples of intricate battles that can have partial solutions without one side or another entirely winning everything.

Let's hypothesize a theoretical partitioning of go endgame positions into tiles with ko and tiles without ko that have omitted moves in their game tree because of bypassing reversible moves. Let's hypothesize that the ko tiles are sorted in some order corresponding to the chilled value of the other tiles. *The best play game trees for the non-ko tiles will need to include all branches of otherwise bypassed moves as long as winning any ko has a larger value.* As kos are settled, the game trees may be simplified to the extent that existing ko positions are no longer threats. We speculate that each node need not be aware of the magnitude of ko positions all over the board, but clearly, having all reversible moves stored in a tile waiting for news of a ko fight being settled runs counter to the spirit of Mathematical Go theory.

This chapter discussed how the difference in the ko rule between Mathematical Go and other rule sets impacts on the correlation of outcomes between MGO and JO/CO. The next chapter will deal with correlating a subset of Mathematical Go for positions too large to be handled exactly by the complete Mathematical Go rule set.

CHAPTER 9. Sente and Gote

This chapter will discuss sente and gote: first from the WW viewpoint, and then from a go player's perspective. Chapter 7 had the beginning of a quote by Wolfe about the instructional value of Mathematical Go which continues [WD, p.1]:

> Sente (the worth of keeping the initiative) and gote (giving up initiative) are issues which are only vaguely understood by all but the finest of go players. This is in part because the concepts are subtle, but also because what is meant by initiative is, by nature, vague and amorphous. The game theory does away with these concepts by providing clearer and more concrete methods, and these in turn will give the go player a better understanding of sente and gote.

Sente literally means "upper hand," but it refers to maintaining the initiative. It is often interpreted inexactly as forcing a response within a local situation. Gote literally means "lower hand," but it refers to giving up the initiative. It is often interpreted inexactly as not forcing a response within a local situation. In WW [BE1, pp. 156-157], we see the theory of sente from the forcing/non-forcing viewpoint. WW is talking about the best play in multiple copies of the same game. Specifically, these are hot games with each option a hot game. Such games are classified as either equitable or excitable. WW defines these two terms as follows:

> A move in an equitable component is usually followed by a reply in a different component, but a move in an excitable one usually requires a response in the same component. A very excitable move poses a grave threat which must be answered immediately; an equitable move does not. In the language of the Japanese game of go:
>
> Excitable moves keep *sente*. Equitable ones don't.

WW analyzes one class of games, where $G = \{ a \mid b \parallel c \mid d \}$, $a \geq b \geq c \geq d$, and $s = a +$

$b + c + d$. If $s > 4b$ or if $4c > s$ then G is an excitable game. Let's play a representative excitable game, H,

$H = G + G + G + G$ where:

$a = 7$, $b = 3$, $c = 2$, and $d = 1$.

$G = \{ B \mid W \}$, or $\{ 7 \mid 3 \parallel 2 \mid 1 \}$.

$B = \{ 7 \mid 3 \}$, Black's move from G

$W = \{ 2 \mid 1 \}$, White's move from G

H is played out four times below. In the left box, White will respond to Black's play. He believes B is sente, forcing an immediate response. In the right box, White will ignore Black's move. He believes B is gote. White's two strategies are tested for either player moving first.

BLACK PLAYS FIRST

WHITE B>G				WHITE G>B				Move
G	G	G	G	G	G	G	G	0
B	G	G	G	B	G	G	G	B1
3	G	G	G	B	W	G	G	W2
3	B	G	G	7	W	G	G	B3
3	3	G	G	7	W	W	G	W4
3	3	B	G	7	W	W	B	B5
3	3	3	G	7	W	W	3	W6
3	3	3	B	7	2	W	3	B7
3	3	3	3	7	2	1	3	W8

WHITE PLAYS FIRST

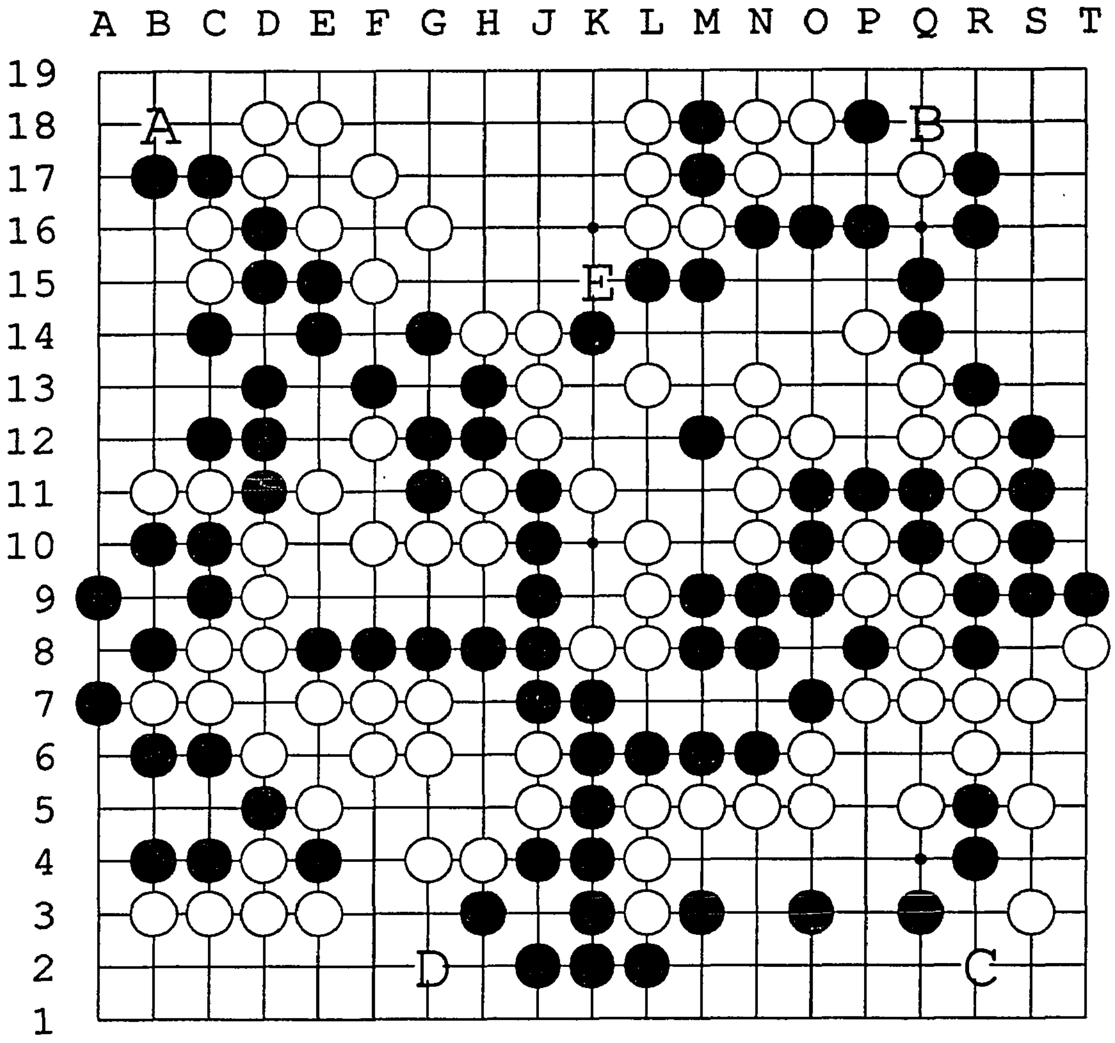
WHITE B>G				WHITE G>B				Move
G	G	G	G	G	G	G	G	0
W	G	G	G	W	G	G	G	W1
W	B	G	G	W	B	G	G	B2
W	3	G	G	W	B	W	G	W3
W	3	B	G	W	7	W	G	B4
W	3	3	G	W	7	W	W	W5
W	3	3	B	2	7	W	W	B6
W	3	3	3	2	7	1	W	W7
2	3	3	3	2	7	1	2	B8

The bottom line indicates the stopping points in H, where all further play is from a number. The value of H in each case is the sum of the four tiles. With either player moving first, White loses a point by not responding in the same component. This is a game with $s > 4b$ which gives a sente move to Black. When $s < 4c$ then White will have a sente move. These are WW games with values that can't arise in go, but they illustrate the sente principle.

The Ishi Press [IP] publishes most of the serious go books in the English language. Additionally, they publish Go World, which is the best English language periodical on go. The author of this thesis searched through these sources for the closest description of Mathematical Go type thinking applied in a sente/gote situation. We found this example appearing in Sakata Eio's "The Middle Game of Go" [SE, pp. 103-113]. Figure 9.1 is from the second game of the 1963 Meijin tournament between Fujisawa Shuko (Black) and Sakata Eio (White). The points A through E are locally optimal playing points. The question is to find the right sequence of areas to play in, given that the best play within the local areas can be determined.

Sakata begins with the classical first approximation. For each local area, determine the value of that area assuming Black moves first, then do the same assuming White moves first. Take the net difference and call it the value of the local area. Within each area, many approximations have to be made. Branches that end in gote are deleted if they have small values, but kept in the calculation if they have large values. Sakata treats values as small if they occur early in the game tree but not after several moves. He also looks into the miai situations across the entire board after a certain branch is taken locally. He evaluates the local values as follows: $A = 12$, $B = 14$, $C = 18\frac{1}{2}$, $D = 8$, and $E = 10$.

In treatments of sente and gote for beginners, the standard formula is that sente is locally forcing and gote is not locally forcing so one counts excess moves within a local area as a guide to the magnitude of a play. Sente is valued as twice gote. Making a gote play before an


A B C D E F G H J K L M N O P Q R S T
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
A
B
C
D
E

White to play, find the biggest endgame point.
From Sakata Eio, "The Middle Game of Go." [SE, p. 103]

Figure 9.1

opponent can play there is also treated sometimes as sente. Sakata takes a more sophisticated approach and discusses in detail the changing situation within a local game tree as other trees are played out. Based on sente over the entire game tree in E, rather than sente considerations that occur in branches of the other moves that have been calculated as higher valued, Sakata concludes that E is the biggest endgame point.

Chapter 6 showed that incentives are often asymmetrical in hot games. In composing a problem, to make choices closer in value, one includes the negative of a game. This concept of miai, balance of choices, helps to explain whole board sente. Within a given area, a move may look like absolute sente, but if the corresponding negative exists on the board, a player can follow the Tweedledum Tweedledee strategy. Beginning go players often will meekly respond to every potential sente threat with the local best response and lose points on every move. (Recall what happens in Figure 6.2 if Black answers White locally after each move!) It's far better to find the negative of the opponent's sente threat, (a sente threat of one's own with equal value), and play there.

Chapter 6 discussed thermostrat. Thermostrat deals with stopping points, obtaining the global mean value, discovering the ambient temperature, and selecting the tile to play in at the ambient temperature. The author of this thesis does not know why plays at A through E in Figure 9.1 are best play, nor does he understand why the local sub-games mentioned in Sakata's analysis are best play. However, the author of this thesis believes that there is a large uncharted area for analytically minded go players to draw similarities between endgame analysis such as discussed by Sakata and the thermostrat algorithm. This analysis of sente/gote would require expertise in both WW methods and go endgames. The Mathematical Go approach would have to be weakened, and the go player's thinking out loud would have to be expanded.

This chapter suggested that in addition to Mathematical Go's impact on whole board endgame problems with small incentives, other WW strategies may impact on hotter go games. The next chapter will discuss computer go, where the tasks to solve are not subject to the nearly exact analysis we have been discussing for the last nine chapters.

CHAPTER 10. Computer Go

This chapter focuses on the direction of computer go. The first working computer go program (perhaps around 30 kyu?) was written by L. A. Zobrist [ZA] in 1970. Since then, computer go programming has become quite popular and programs have gotten stronger. Bruce Wilcox popularized computer go among go players with a series of articles in the Journal of the American Go Association [WB]. Anders Kierulf [KA] created a user-friendly interface for go developers, the Smart Game Board, which was used by three of the six contestants in the 1991 North American Computer Go Championships. David Erbach [ED] edits and publishes Computer Go, a quarterly journal entirely devoted to this field. Adrian Mariano [MA] is the system administrator for the INTERNET bulletin board rec.games.go which includes many articles on computer go along with traditional go articles. Mariano also maintains the archives on the anonymous FTP site milton.u.washington.edu, in the directory public/go. These archives include a concatenation of several computer go bibliographies that is 30 pages long. Wolfe's toolkit [WD2] is located here, along with many other interesting files. The computerization of go has long been an inherently happy avocation. National and international computer go tournaments are regularly conducted. Perhaps Ing Chang-Ki's $1,600,000 prize (see Chapter 3) has stimulated interest.

The current level of championship programs have been estimated to play in the 8-15 kyu range, but, as David Fotland [FD] explains:

> Each program may have certain anomalous vulnerabilities which can be discovered after a few games and subsequently exploited....
>
> None of the current programs can learn from their own mistakes; when the same situation comes up they will make the same bad move again....

> I know someone who was having trouble beating Many Faces [of Go, Fotland's championship program] at 13 stones until I suggested he could beat it at 29 stones. He spent a few weeks trying odd moves and found some weakness, and now he has no trouble beating it at 29 stones....
>
> Each of the programs has different weaknesses, but they all tend to collapse tactically in a complicated position, so if you attach and crosscut a lot you can usually win big.

This inability of computers to learn from experience, once their weaknesses have been exposed, is to be contrasted with human go players. Boon's Goliath go program won the 1991 World Computer Go Championship and the right to challenge three humans with Boon's program getting a 17 stone handicap every game (Chinese rules). Boon's program won all three games, a prize of $8,000, and the right for the next challenge at 15 stones. A report of the event [SL] says the three humans analyzed their first round results and developed better strategy by the next day. Boon lost all three second round games. One could speculate whether after six games acquaintance with Goliath, these human players would ever again lose at 17 stones. One could also speculate on how much time and effort would be needed for a similar improvement by any go program.

Computer game playing programs have been constructed for many games with various degrees of success. Let's compare checkers, chess, and go. In 1970, Zobrist provided a rough estimate of the number of nodes needed for a winning game tree based on the number of moves in a typical game and the average number of reasonable choices per move. His figures were:

Checkers: 3×10^{20}

Chess: 5×10^{50}

Go: 1×10^{100}

In 1991, the number of checker nodes was better approximated by actually counting them. A checker program (built by a team of Canadian programmers) was better than all human checker players except one. Not only that, the entire best play game tree was nearly completed. As the team reported in SIGART [SJ, pp.3-5]:

> There are $5 * 10^{20}$ possible checkers positions. To solve the game of checkers, determining the game theoretic value of the game requires a proof tree of roughly 10^{11} positions. The problem is knowing which 10^{11} positions! With today's technology, checkers can be solved, given sufficient computational resources. Consequently, fairly soon (within a decade) it may be possible to create a "perfect" checkers player.

Their approach involved state-of-the-art database technology to store and retrieve almost all positions. The late and mid endgame were exhaustively played out. Beginning with storing published master games, the team added more and more exact best play game trees to their growing database.

Chess programs have now reached grandmaster strength. Their games are studied by players just like the games of other masters. "What would Deep Thought (the current computer chess champion) play?" is as valid a question as "What would grandmaster John Doe play?"

These programs for checkers and chess are heavily dependent on specialized hardware and vast search trees. Someone has noted that for several years in the 1980's, each half-ply increase in search for a computer chess program was worth 100 rating points. From this, one might deduce that size of the search tree tracked more closely to playing strength than any heuristics or other playing strategy.

Various methodologies have been proposed for use in computer go programs, among them, cellular automata theory, neural nets, machine learning, syntactic pattern recognition, supercomputers, the connection machine, and other parallel processing. Wolfe adds Mathematical Go to the list [WD1, p. 2]: "A good program will require a strong understanding of how to integrate what is known about nearly separated positions." We disagree. Recall the quote from Wolfe in Chapter 7 [WD1, p.1]: "Many of the positions analyzed do not tend to come up in contexts which take full advantage of the subtleties of the results." For a computer program (or human) to successfully navigate its way through the opening, midgame, and early and middle stages of the endgame; to stay within a point of its opponent, then lose the game because of failure to use Mathematical Go theory is quite unlikely. *Wolfe has already written the toolkit to handle most situations* [WD2]. The interface to Wolfe's toolkit should be easy to program when compared with other methods proposed for improving computer go. However, for the near term, there are more pressing issues.

The hardware, software, system, and management techniques that build checker and chess programs which play better than almost everyone in the world are fully available to the computer go community. Yet, computer go programs play on the level of a weak novice. Apparently the difference in the effectiveness of existing techniques depends heavily on the size of the solution space. Most go programmers feel Zobrist's estimate of 1×10^{100} is much too low; in fact, figures as high as 1×10^{200} have been suggested. Checkers is well suited for database storage because as the game draws to a close, the number of pieces and potential moves diminishes, and the number of moves for the complete tree is within the reach of existing computational resources. Like checkers, chess is amenable to exhaustive enumeration because as the game draws to a close, the number of pieces and potential moves diminishes. There exist several end game positions, such as King, Queen, and Pawn versus King and Queen, that have been solved by computer but cannot be solved by any human within normal playing time. However, because chess has more moves than checkers, a complete solution by enumeration is still impractical.

The nature of go is to have more plausible moves as the game progresses because most pieces are not removed from the board. Mathematical Go has pointed the way to a new form of tree that handles search trees without strict alternation. It has also shown how local areas can be cancelled out in computing board-wide scores. (This miai method of counting is often used by go players during a game.) Both Wilcox [WB] and Fotland [FD] have written how local area database techniques are used in the go opening, and how tacticians do search in the endgame.

In August 1991, the author of this thesis conducted interviews with most of the contestants in the North American Computer Go Championships held at Rochester, N.Y. We shall refer to those interviewed as "the panel" [PA]. We solicited the panel for their opinion on what would be required to strengthen the existing computer go programs participating in the championships.

The panel was unanimous on two items: theme-based search is the only profitable way to program for the next 10 years, and the most important advances in computer go can only come through acquisition of more "go knowledge." The panel agreed that the state of the art techniques used in other game-playing programs are available to the panel, but would not be useful until advances in search and understanding human go knowledge were made.

Theme-based search is Wilcox's label for what other panel members described with different terms. In searching a tree, the most difficult part of any algorithm is to decide if and when to open a node. Next, when one opens a node, there is the question of to what depth one should continue exploration. Finally, how does one evaluate the subtree that has been opened? One creates theme experts or move generators to search for individual themes. The input for detecting a theme comes from the pattern recognition portion of the program. Some examples of themes known to go players are accumulated in Go Proverbs Illustrated by Kensagu [KS]: "hane at the head of two stones," "make ponnuki," "consider the 1-2 point," and so forth. Without going into needless detail, we note that a pattern recognizer looks for precursors to known efficient shapes and hands over this information to the theme expert/move generator for further processing. The theme expert typically has an extremely abbreviated search tree accompanied by a theme-specific evaluation scheme to quickly assess the results of a local situation. We have an example of theme-based search from the human point of view in DeGroot's "Thought and Choice in Chess" [GA]. DeGroot, a psychologist and strong chess player, devised a number of experiments with flashing screens and other equipment to assess the difference in search depth among chess players of varied strengths. He concluded from his experiments that weak, intermediate, and strong players, in almost all situations, have internal search trees of the same breadth and the same depth. The difference among them was expertise in which local area to look at and within a local area, which move to select. A strong player might consider 5 nodes to explore that were totally disjoint from a weaker player's selection. The selection process came from better recognition of certain piece configurations that were considered cooperative.

After the theme experts/move generators have assessed the local phenomena, global evaluation may occur. Boon's Goliath is said to be two stones stronger than any other program because Boon began working on global lookahead two years before anyone else. Global lookahead tries to resolve all the captures over the board until each group is quiescent, a very time intensive computation. Chen notes that in quiescent conditions, a program should ignore global considerations; e.g., always connect against a peep to avoid squandering time over the alternatives.

Because any program is as weak as its weakest link, a program needs a strong tactician to handle the evaluation of life and death problems. Fotland's Many Faces of Go went undefeated in the 1991 tournament, but in the individual games, the program was not always ahead from start to finish. When Many Faces of Go played against Chen's Go Intellect, Go Intellect mapped out many areas of near territory in the middle game that should have been sufficient for an easy win if a conservative human of weaker strength had taken over. But Go Intellect kept on trying to garner more territory when it should have been consolidating, and Fotland's superior tactician killed a number of Chen's groups and won. Fotland has given an excellent summary of how his program works [FD].

The panel suggested parallel processing might be used in computer go programming in several ways. They felt the best use would be for each independent processor to look for a totally independent theme on the board and report its presence to a central controller. The idea of agreement being reached among processors by voting protocols was thought of as absurd. Another suggestion was to network all the contestants into one central control unit. The panel felt that an optimistic upper bound for this combination of existing programs would be one kyu better. One member joked that it would probably play two kyu worse.

Checkers is an excellent example where tuning of algorithms based on feedback from a database of games led to a huge success [SA]. The panel felt that tuning, even with training runs on a supercomputer, would be worthless or positively harmful. Tuning works against the existing evaluation functions in a program, and could not be helpful without the injection of go knowledge. One member felt that having a Cray available to tune all the algorithms and evaluation functions for an extended period of time might improve a program by $\frac{1}{3}$ kyu. For comparison, humans playing at the same level as top go programs (around 12 kyu), frequently improve by 2 to 3 kyu after playing 50 additional games, *even without study*.

The panel felt confident that the computing resources they were using were adequate for program development. They all stated that there was a clear need for incorporating go knowledge into their programs. Fotland had stated that he thought a very useful contribution to Many Faces of Go was being made by his getting regular go lessons from a stronger go player.

It is interesting to consider parallels with the field of expert systems. There are many cases where the knowledge of a field has been incorporated into an expert system program so successfully that the program became better at some tasks than the human who had originally done the job. On the other hand, there are many cases where the expert system approach failed miserably. Just about every strong computer go program has been written by a dan-level player. This means that the expert and the expert system developer are one and the same person. The identification and utilization of "go knowledge," the study of how strong go players think, is a hot open question.

Epilogue

In all previous chapters, the author of this thesis has cited the appropriate references. This chapter offers his own ideas based on memories where the references mentioned may be more ambiguous, less precise.

Why is this thesis named "Mathematical Go: An Analysis"? Casti, in "Paradigms Lost," explained the difference between modeling and analysis. Modeling abstracts from a real world system to an essential model that explains observations and is useful for predictions. The model is smaller than reality. Analysis maps from a real world system onto several other fields; it makes more associations and connections to other systems; its output is larger, and more confusing, than the original system.

Mathematical Go, unlike the physical sciences, deals in the realm of perfect information. Kuhn says scientists like to work at solving puzzles. Kuhn also says that for the history of science to be accurate it must understand the paradigm of the day and how it shifts. On a small scale, Berlekamp's and Wolfe's work has shifted the paradigm of cognitive science. As stated earlier, Mathematical Go is a breakthrough in psychology.

Philosophers foreshadow science, and scientists foreshadow the mind-set of the population at large. Wittgenstein tries to place ordinary language on a scientific basis. Chomsky focused on the underlying grammar. Lenat is working on establishing a database for the common sense use of words. Mathematical Go is an example of describing non-ordinary, counter-intuitive notions in ordinary language.

The future of Mathematical Go can extend in three directions, inside out, outside in, and in tandem with cognitive science.

Extending Mathematical Go inside out means adding more facts and theories to existing rules. Things to consider are: specifying values for shapes other than corridors, extending the analysis of rooms to larger tiles, exploring sockets where the connection is not absolute, evaluating ko with fuzzy logic to account for the different objectives of playing to win or draw, characterizing multiple ko tiles, lessening the precision of results to sum tiles of higher temperature, and so forth.

Extending Mathematical Go outside in means applying the mathematical theories of the very late endgame to other areas of go, to other games, and to other information systems. Chapter 9 suggested that sente and gote considerations in the late endgame may be better understood with a thermostrat-like algorithm. Other loopy games may share the results of Mathematical Go. Perhaps some real physical systems can be modeled with loopy game trees. The Mathematical Go issue of when a tile becomes immortal is confused with the go issue of life and death.

Extending Mathematical Go in tandem with cognitive science means using the control structure and transformations of Mathematical Go to model how go players think, and using new probes into how go players think as directions for how to program go.

Go players think in many ways. They compute with their left-brain. They develop esthetic values with their right brain. Localized losses in stroke victims who play go show that opening knowledge is stored in the right brain, endgame ability is stored in the left brain.

Is there some relation between whether a person learns to read from ideographs (versus alphabet characters) and how he learns to play go? Why do the peoples of Korea, Japan, and China have such a high prevalence of strong go players compared with the occident? Herb Doughty was the manager of the Berkeley Go Club for many years and taught hundreds of children. He observed that children who learn to play only by patterns and not by any verbal constructs will eventually attain higher rankings than their peers who learn themes specified in words injected upon them by unthinking trainers, although in the short run they fare worse. Does this relate to the concept of pattern recognition versus theme-based search?

Why did professional go players with prodigious memories and calculating prowess not beat Berlekamp? What connections are they missing?

Reitman tried to analyze the thinking of Kerwin (now a one dan professional) while Kerwin played go. This was a brief experiment. Is there some psychologist who will do for go what DeGroot has done for chess?

Would a high dan level computer go program, which may take many years to develop, satisfy the Turing test as a thinking machine?

Where are the higher levels of thought carried out in the human? Penrose says mathematicians think in a special way. He rails against the strong AI hypothesis. Whorf says the higher levels of thought are carried out only in language. Could Mathematical Go convince us of an alternative formulation? Is visualization of patterns the nexus of the arithmetic logical unit of the brain, with words serving only to synchronize Minsky's society of processes in our mind.

Go teaches calmness and tranquility. Is there some endorphin-like like release of an equanimity-increasing factor that is stimulated by concentrating on non-language themes?

Descartes feared endless recursion, so he named a God to break out of the loop. Will familiarity with Mathematical Go's recursions without basis help cure the fear of recursion in other mathematicians?

Cantor introduced a hierarchy of infinities where only one had sufficed. What will be the impact of "Winning Ways" hierarchy of infinitesimals?

Luria described the mind of the mnemonicist. Lorayne unraveled the tricks of memory for the ordinary person. Are these related to the go study performed by an "insei" (aspiring go professional)?

Lorenz showed us several ways to look behind the mirror, to see who we are and where we came from. The signature of one contributer to rec.games.go is "Life is a metaphor for GO!" Go is large enough to be distributed throughout the brain. Go moves are made by go players who think with human brains. Mathematical Go theories are made by mathematicians who think with human brains. Computer go programs are written by computer specialists who think with human brains.

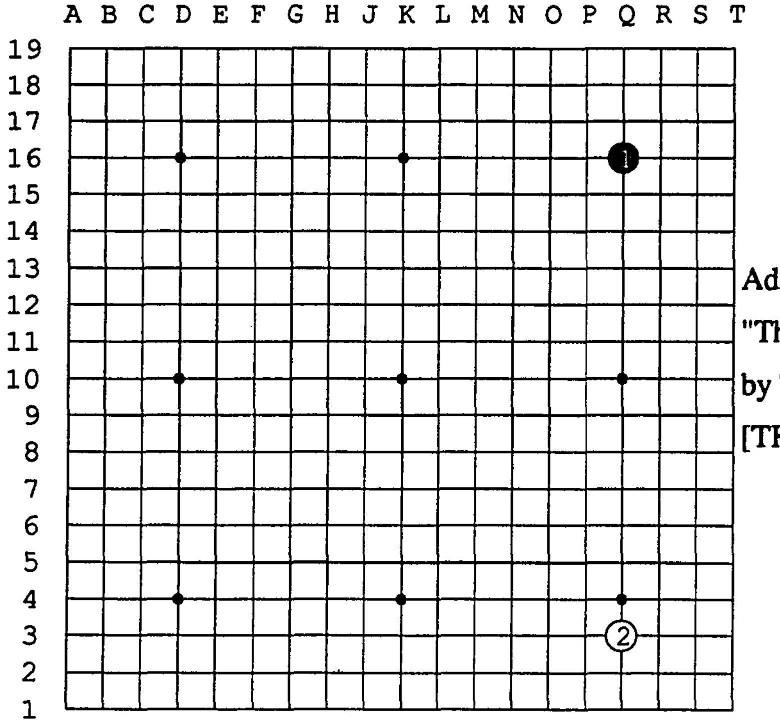
Using this triumvirate of cultures should enable us to better answere the questions, who we are, where we came from, and where are we go-ing. Go is learning. Go is fun. The three go cultures are a window into the mind of man.

APPENDIX. Two examples of influence.

These two examples of go positions will contrast go with Mathematical Go. The first example shows influence without immediate profit. The second example shows influence with sacrifice of immediate profit.

Figure A.1 is adapted from "The Direction of Play" by Takeo Kajiwara [TK, p.57]. Black plays first at Q16, a 4-4 point with the corner counting as the 1-1 point. White plays second at Q3, the 3-4 point. The 3-4 point in relation to a corner, irrespective of symmetry, is the most common move made in the opening. However, in relation to the previously played Black 1, Kajiwara concludes it "is probably a losing move." He supports this with several pages of analysis. The title of the book aptly describes his thesis that the shape and location of stones in one region of the board exude power and have influence in specific directions that should guide subsequent play. Mathematical Go has no such concept.

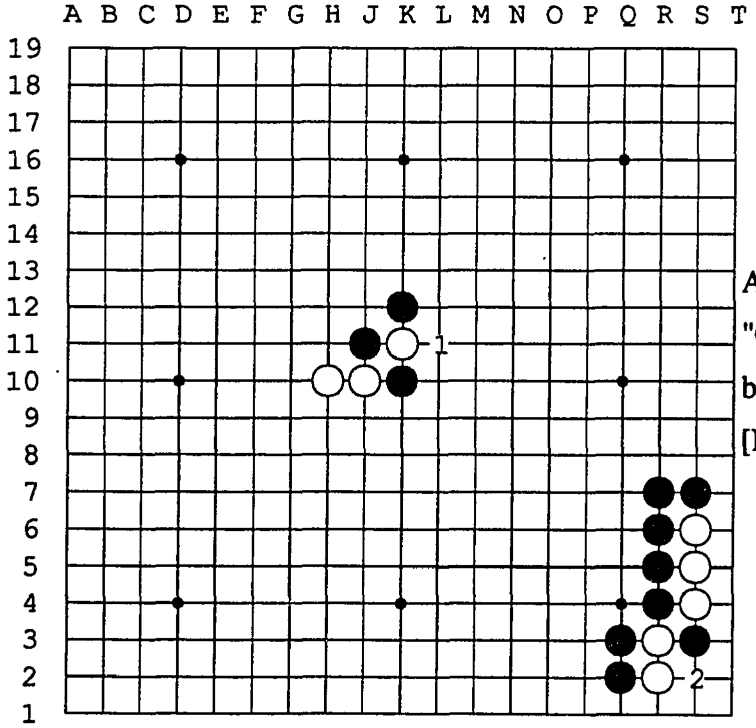
To a novice, high kyu player, the above comments would be difficult to understand, even after reading Takeo's analysis. A low kyu player, after going through Takeo's analysis, might lend some credence to his conclusion. A dan player might understand the concept without benefit of Takeo's arguments. The point of this example is to focus on one idea: go, especially in the opening, is a whole board game. Many books on the opening focus on joseki, or how to play fixed established patterns locally in the corner. The books always add that one must view joseki only in relation to the entire board.



A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
1
2

Adapted from
"The Direction of Play"
by Takeo Kajiwara,
[TK, p.57]

Figure A.1



A B C D E F G H J K L M N O P Q R S T
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
1
2

Adapted from
"Go Proverbs Illustrated"
by Kensague Degoe,
[KS, p. 221]

Figure A.2

Figure A.2 is adapted from Go Proverbs Illustrated by Kensagu Segoe [KS, p. 221]. The problem is which is the best play for Black, at 1 (L11) or at 2 (S2). The position is contrived but illustrative.

If Black plays at 1, he captures one prisoner, forms a shape that is very hard to kill and won't be separated by a White play at 1 into three weak stones. White should reply at 2, capturing one prisoner (inevitably), forming a live group worth around six points, and having the potential to extend at Q1 and T7. If Black plays at 2, he captures five prisoners (inevitably), and gains thirteen points of secure territory. White should reply at 1, separating Black's three stones into two weak groups.

The author concludes that the difference in who gets to play first at 2 is worth around 30 points, and that playing first at 1 is clearly superior. Here, influence prevails over clear profit. Mathematical Go has no such concept.

Bibliography

[AGA] American Go Association, "AGA Tournament Rules," P.O.Box 397, Old Chelsea Station, New York, N.Y. 10113, 1991.

[BE1] Elwyn R. Berlekamp, John H. Conway, Richard K. Guy, "Winning Ways for your mathematical plays," Volumes I and II. Academic Press, London, 1982.

[BE2] Elwyn R. Berlekamp, "Introductory Overview of Mathematical Go Endgames," Proceedings of Symposia in Applied Mathematics, Volume 43, American Mathematical Society, 1991, pp. 73-100.

[BE3] Elwyn R. Berlekamp, personal communication.

[BM] Mark Boon, "A Pattern Matcher for Goliath," Computer Go, Number 13, Winter 1989-1990, pp. 12-23.

[CJ] John H. Conway, "On Numbers and Games," Academic Press, London,1976.

[CR] Raymond Chen, personal communication.

[DJ] James Davies, "The Japanese Rules of Go," Go World, #12, March 1979, pp 43-46.

[ED] Erbach, David, editor and publisher "Computer Go," Winnipeg, Canada, 1986 - present.

[GA] A. D. DeGroot, "Thought and Choice in Chess,". Mouton, The Hague, 1965.

[FD] David Fotland, in [RGG], Documentation for "The Many Faces of Go," 1991.

[GW] "Go World," published quarterly by Ishi Press International, #1-#66, 1976 - present.

[IP] Ishi Press International, catalog, San Jose, California.

[IT] Ikeda Toshio, "The Rules of Go," Fujitsu, Tokyo, 1992.

[KA] Anders Kierulf, Ken Chen, and Jurg Nievergelt, "Smart Game Board and Go Explorer; A study in software and knowledge engineering," Communications of the ACM, Volume 33, #2, February 1990, pp. 152-166.

[KS] Kensagu Segoe, "Go Proverbs Illustrated," Japan Go Association, Tokyo, 1960.

[KY] Kano Yoshinori, "Graded Go Problems for Beginners, Volume 4," Nihon Ki-in, Tokyo, 1990.

[LA] Arthur Leff, personal communication.

[LE] Edward Lasker, "Go and Go Moku," Dover, New York, 1934.

[MA] Adrian Mariano, [RGG] system administrator, and milton.u.washington.edu public/go archivist (an anonymous FTP site).

[MD] David A. Mechener, "A Matter of Life and Death," The American Go Journal, Volume 25, Number 4, Fall 1991, pp 30-31.

[NR] Richard J. Nowakowski, "..., Welter's Game, Sylver Coinage, Dots-and-Boxes, ...," Proceedings of Symposia in Applied Mathematics, Volume 43, American Mathematical Society, 1991, pp 155-182.

[NY] Nagahara Yoshiaki, "Strategic Concepts of Go," Ishi Press International, San Jose, 1972.

[PA] Lynn Bues, Ken Chen, David Fotland, Anders Kierulf, Martin Mueller, Bruce Wilcox. Personal panel communication, August 1991.

[RGG] INTERNET Bulletin Board, rec.games.go.

[SA] A. L. Samuel, "Some studies in machine learning using the game of Checkers. II - Recent progress." IBM Journal of Research and Development, 11, No.6, 1967, pp. 601-618.

[SE] Sakata Eio, "The Middle Game of Go," Ishi Press International, San Jose, 1971.

[SJ] Jonathan Schaeffer, Joseph Culbertson, Norman treolar, Brent Knight, Paul Lu, Duane Szafran, "Checkers Program to Challange for World Championship," SIGART Bulletin, ACM Press, April 1991.

[SL] Shao Lejun, "The 7th International Computer Go Congress Championship," Computer Go, Number 16, Spring/Winter 1991, pp. 8-9.

[TK] Takeo Kajiwara, "The Direction of Play," Ishi Press International, San Jose, 1979.

[VJ] John Von Neumann and Oskar Morgenstern, "Theory of Games and Economic Behavior," Princeton University Press, Princeton, 1944.

[WB] Bruce Wilcox, instant Go and Computer Go articles, originally published in the AGA Journal, available from Bruce Wilcox.

[WD1] David S. Wolfe, "Mathematics of Go: Chilling Corridors," Doctoral Thesis, University of California at Berkeley, 1991.

[WD2] David S. Wolfe, archived in [MA], theory.sh.Z

[ZA] A. L. Zobrist, "Feature Extraction and Representation for Pattern Recognition and the Game of Go, " Doctoral Thesis, The University of Wisconsin, 1970.