Diplomacy — A.I.

Alan Ritchie

A dissertation submitted in part fulfilment of the requirement
of the Degree of MSc in Information Technology
at The University of Glasgow.

September 2003

Acknowledgements

I would like to thank Dr Ron Poet for his help and advice on C++ and the game of Diplomacy.

I would also like to thank Ally, Bob, Claire, Doug, Fuzzi, Graeme, Karen, Neil, Omar, Sandy, Sandra, Tam, Tim and Tom for their time spent playing Diplomacy and other games of imperfect information, as we figured out why I lost, how to win, and what the program should try. And finally my thanks to Lindsey for her time and patience, despite having no interest in the game.

Abstract

Diplomacy involves seven players negotiating and fighting across Europe as they attempt to conquer the continent. It is an entirely deterministic game, but players move simultaneously producing a game tree too large to be searched by normal methods.

Traditional search methods struggle with imperfect information, and fail with simultaneous moves.

This report describes a program that plays a simplified five player no-press variant of the game at a novice level, being moderate tactically but weak at the strategic level.

It details a simplified 2 player and 5 player version of Diplomacy, which the program was tested against, and examines in depth the moves made compared to what moves would be expected from a human player. The program does not implement fleets.

Contents

Chapter 1

Introduction

Diplomacy is a seven-player board game set at the start of the 20th century. Each player controls one of the Great Powers of Europe, and play is on a board representing a map of Europe. The object of the game is to gain control of Europe.

There are three features that make Diplomacy different to other games. Firstly, the game is entirely deterministic, with clashes between powers resolved by weight of numbers rather than dice or another random factor. As each power is equally strong at the start, no single power can gain an advantage over their rivals without the aid of an ally.

The second feature is that the game includes seven players rather than two. Most research concentrates on two player games, but in Diplomacy there are additional problems like deciding when to cooperate with opponents. A possible move for one player must be put in context of how all the opponents see it, whereas normally a good move is automatically bad for the opponent, and vice versa.

Finally all players move simultaneously. Almost all other games are played in turns. This makes searches of the gamespace very difficult because players do not have perfect information, and cannot base decisions on probable outcomes because they cannot calculate the necessary probabilities with any reliability.

Microprose released a commercial implementation of the game in 1999, but it was criticised for having very poor AI [1] which made the game very easy to beat.

SeaNail is the most advanced available alternative. It features a GUI and powers that actively negotiate with each other, but do not change their views toward each other. It does not currently appear to be being actively developed.

The Diplomacy AI Development Environment (DAIDE) is a server and adjudicator designed to enable play between competing agents. It has a limited press syntax to allow communication between players and runs over TCP/IP. The rationale is to allow several bots to be designed by separate groups and play over a common framework. A number of bots are listed as under development, but none are listed as playing much better than random moves and all are limited to no press games. DAIDE specifically prohibits attempts to communicate outwith its communications protocol.

Danny Loeb worked on the theory of multiplayer games [2], and his Diplomacy Programming Project produced the negotiation protocol that DAIDE is based upon. A diplomat produced by Kraus and Lehmann [3] was based around a negotiating agent rather than the game strategy, but negotiated like a human player. However neither of these two programs are widely available.

Diplomacy is suited to play by email, having been played by mail since the 1960s. A program called the Judge has been developed to automatically adjudicate email games,

and a wide range of variants are supported.

This program aims to play no-press Diplomacy, at a modest level. It is specifically designed to work with the judge and play email games, but to play through a human who will handle the emails to and from the judge.

The program was tested on a two player and five player map, which was entirely land. The two player game was played quite strongly, but in the five player game struggled to concentrate force against a single enemy, and was unable to achieve a victory.

The program cannot handle the entire game of diplomacy, as fleets cannot convoy armies across bodies of water, do not specify which coast they wish to move to when moving to bicoastal provinces, and are not built in adjustment phases.

Chapter 2 is an overview of the game rules and Chapter 3 briefly discusses strategic and tactical theory. Chapter 4 covers some game theory relating to this project. Chapter 5 is a description of the program. Chapter 6 uses a simple two player version of Diplomacy, and Chapter 7 a five player version, to test the program, with both including some comment on how each player might reasonably approach the first few turns and what mistakes were made. Chapter 8 covers some possible future developments.

Chapter 2

The Game of Diplomacy

This section is a brief summary of the rules of Diplomacy. The full rules [4] can be obtained from the Avalon Hill or Hasbro websites, or by buying the game.

Each player controls one of the Great Powers of Europe: Austria-Hungary, England, France, Germany, Italy, Russia and Turkey; (also referred to as "nations" and "countries") which are represented on the board by their armies and fleets.

The game is played on a board representing Europe, see Figure 2.3. The board is split into provinces, either inland, coastal or sea, which can only be occupied by a single unit (army or fleet) at a time. Thirty-four provinces are supply centres, control of each entitles the owner to a single unit. Each power begins with three centres (apart from Russia which has four) and the rest start as neutral.

Each year is split into Spring and Fall turns, starting with Spring 1901. In each turn each player can negotiate with their rivals, then give orders to all their units. These are simultaneously revealed and processed. All units that were dislodged must now retreat and are given orders which are then revealed and processed. At the end of each Fall turn all occupied centres fall under the control of the player occupying them, and unoccupied centres remain under the control of the last player to have occupied them at the end of a previous fall turn. Then each player counts the number of units and supply centres that they have, and builds or disbands units to ensure that they have no more supply centres than units. Units can only be built in a power's original centres. If one player controls eighteen centres at the end of the Fall, then they win.

Crucially, only one unit can occupy a province at a time, and all units have equal strength. Units can move, hold or support a move of another unit. Supporting effectively transfers the strength of one unit to another, but only if the supporting unit is not attacked by another unit. If the supporting unit is attacked (by a unit that it is not supporting an attack on) then the support has no effect, and is described as being "cut", even if the attacking unit is unit is dislodged. If two or more units contest a space, then the one with the most support occupies it, dislodging the occupants. In figure 2.1 the support from Silesia is not cut, because the attack is coming from the province that the unit is supporting a move against, so the army in Warsaw is dislodged. If both units have equal strength then they standoff and "bounce", and neither moves but any occupants of the space are not dislodged. Figure 2.2 shows a German army in Prussia moving to Warsaw, supported by an army in Silesia. But the army attacking Silesia from Bohemia cuts support, so everyone bounces. Support cannot be refused, even if from a unit of another power and no unit can cause another unit belonging to the same power to be dislodged. Units cannot swap places, unless one is convoyed, but three or more units can

move in a circle.

Figure 2.1: A successful attack, from the Diplomacy Rulebook[4]

Figure 2.2: Cut support causing a standoff, from the Diplomacy Rulebook[4]

Fleets can only move along coastlines or into sea areas. Three provinces — St Petersburg, Spain and Bulgaria — have two separate coast lines, and a fleet on a coastline can only move to map areas joining that coast. Kiel and Constantinople have a single coast, and fleets can move in one side and out the other. Denmark is connected to Sweden, and armies can move between, but Spain and North Africa are not connected.

The winner is the first player to control 18 supply centres. The assumption is that the 16 units belonging to the other players cannot stop the eventual conquest of Europe. If no player wins outright, then all surviving players share equally in the draw, regardless of how many centres each actually held.

The game is suited to play by mail, and now over the internet, with players sending email to a judge program, which will forward messages ("press") to other players, process orders and perform other administrative tasks.

There are many variants of diplomacy, using different maps and/or some additional rules. Of greatest interest is "No Press" diplomacy, which prohibits communications between players (Diplomacy without diplomacy). This allows the program to act purely on the state of the map, and opponents' previous orders. Also of interest is "Limit Press", where players have a strict set of phrases that can be used in negotiations.

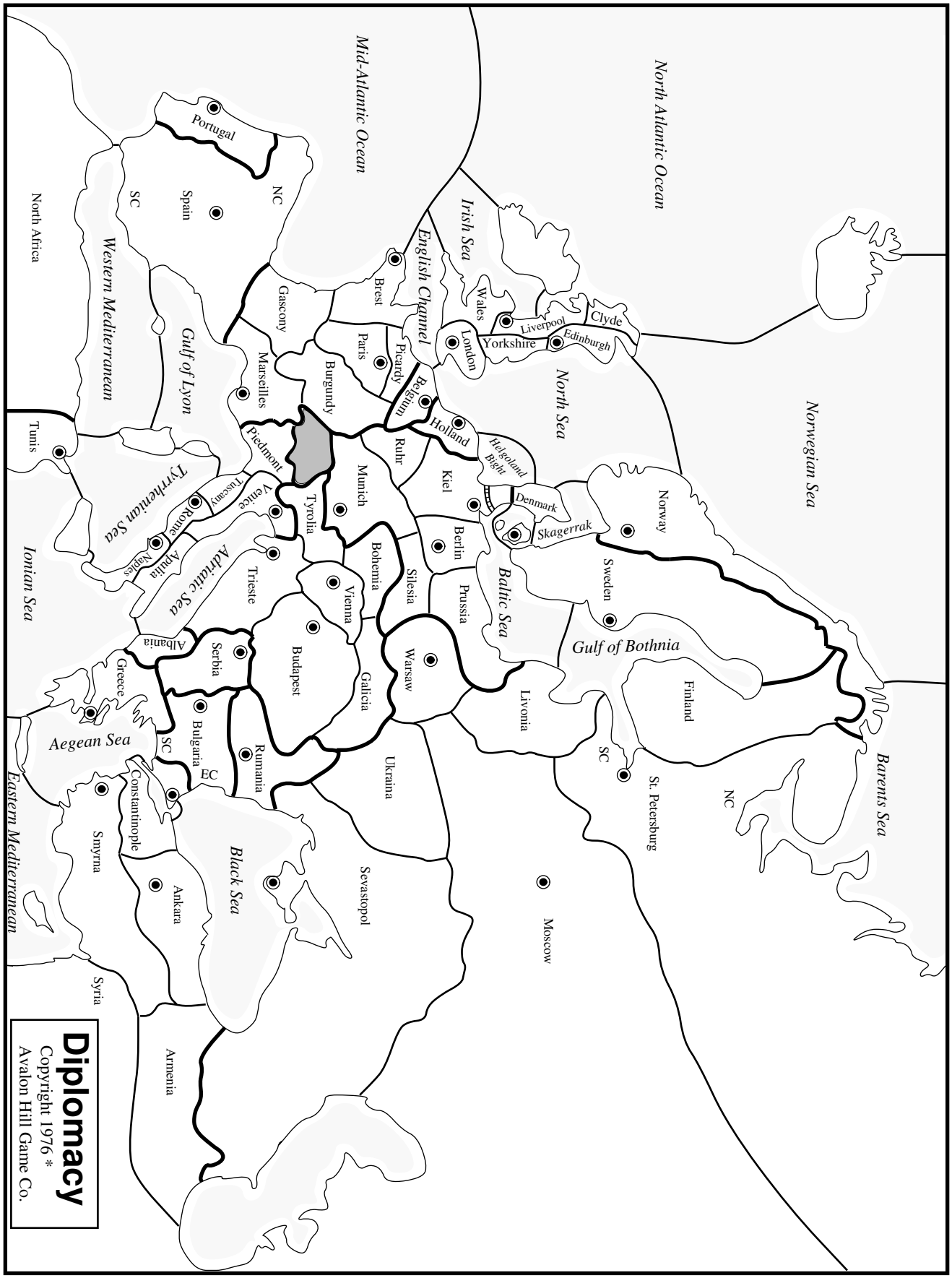

Figure 2.3: The standard map

Chapter 3

Theory of Diplomacy

Diplomacy has a large body of literature, in a similar style to chess, on theory. Many articles were published in the magazines written by postal game masters that were a vehicle for the games that they were running. These and other articles can now be found on the websites supporting the game, including the Diplomatic Pouch at http://www.diplom.org/. Richard Sharp wrote a book[5] including openings for each power, and further advice on face to face and postal play. Most describe a range for opening moves for Spring 1901, examines the probable openings from the opponents, and details possible continuations for Fall 1901, but little further. Some look to the eighteen centre victory condition, work out the eighteen centres that would be most easily captured with a couple of alternatives, and focus on how to start getting there. As postal games and email games can be recorded, there exists statistical information on winners, survivors, when powers are eliminated, and how often players chose particular sets of opening moves. With six other players there is much intrigue as promises are made for Spring 1901 to everyone, and some broken by the first set of moves, and some more by the second. Hence the situation may not fully crystalise before Spring 1902, making planning beyond then difficult. Equally many favour openings that are ambiguous entirely for this reason, as they can be sold as all things to all people, and provide the flexibility to react to events.

One example for England boils down to the fact that England requires to gain access to the Mediterranean or face a difficult overland attack into central Europe. But the entrance to the Mediterranean can easily be held against her. Combined with the difficulties of progressing past St Petersburg after attacking Russia, it suggests that an early all-out attack on France is the only sensible way to go. But Norway is the only guaranteed centre for England in 1901, and if France suspects an attack then England risks losing all chance of a foothold on the continent.

Another source of theory is the endgame. Because Switzerland is neutral, and hence impassable, and a province can only be forced if you have more support than your opponent, there exist "stalemate lines" across the map which can be held indefinitely by the right combination of units. Critically a stalemate line must enclose at least as many centres as units it requires to maintain, and each province must be supported by one less units than could possibly attack it. Successfully holding a stalemate line against an alliance of opposing players will either force a draw, or force the alliance to break up, hopefully allowing you to take advantage of the stab.

At the purely tactical level there are a couple of ploys for relatively common situations. "Self-bouncing" occurs when a player orders that two units move into the same province. As both have equal strength they will both bounce, and stay where they are. But any

enemy unit moving (without support) to any of the three locations will bounce as well. This allows a player to defend three provinces with only two units. But this can easily be countered if an opposing unit supports one of the moves. As the support cannot be refused the supported move will succeed, allowing an opponent to move into the space left behind it. This is the situation in figure 3.1. In Fall 1901 Austria orders both Vienna and Serbia to Budapest, defending Budapest from Russia in Galicia, and keeping a unit in Serbia to claim it in the winter and gain a build. But Russia supports the move from Serbia, instead of attacking Budapest, so Austria does not gain Serbia.

Figure 3.1: Self-bouncing, from the Diplomacy Rulebook[4]

Another is a "beleaguered garrison". A unit in an advanced position may be required to attack, often an attempt to cut support or bounce with approaching enemy reinforcements. But the unit then cannot be supported because it is moving away, so (if it bounces) it is vulnerable to being dislodged by an enemy attack. To stop this the unit can be "attacked", with support, by other units of the same power. The attack cannot dislodge the friendly unit, but is intended to be enough to bounce with an enemy attack, (which does not dislodge a unit in the province that was bounced in) so the unit can cut support without the threat of being dislodged.

In figure 3.2, the Russian army in Berlin attacks Kiel, and the Russian fleet in Skagerrak attacks Denmark, supported by the fleet in the Baltic. But the English fleets stop Russia from taking either centre by using Denmark to attack Kiel, and by supporting the fleet in the North Sea into Denmark. Kiel is not occupied because Berlin and Denmark bounce. And the fleet in Denmark is not affected by the standoff between the supported fleets from the North Sea and Skagerrak.

Figure 3.2: A Beleaguered Garrison, from the Diplomacy Rulebook[4]

Chapter 4

A.I. Theory

4.1 Theory of Games

There is a large body of literature on the theory of games, and more generally theory on decision making. Charles Babbage designed a machine to play Tic-Tac-Toe, and believed that his analytical engine could be programmed to play chess, but neither were built. Von Neumann applied game theory to economic decisions and first described the minimax search. Alan Turing wrote the first chess program for a computer, but it was never run. But in recent years Gary Kasparov and then Vladimir Kramnik, both Chess World Champions, have played exhibition matches against computers, and both struggled, with Kasparov losing $3\frac{1}{2}$ to $2\frac{1}{2}$ to Deep Blue, and Kramnik drawing 4-4 against Deep Fritz.

Games can be classified in different ways: as perfect or imperfect information, as games of chance, if draws are possible or not, zero-sum or not, and so on. Chess and draughts are both perfect information games. Backgammon and Risk are perfect information games, but games of chance with the outcome depending on dice. Card games such as Bridge are imperfect information games, but not games of chance because when play begins the cards are already dealt. However players often resort to probable outcomes when deciding how to play a hand. Poker is considered an imperfect information game of chance, because the strengths of a hand will change depending on what cards are dealt during play. However it is equally accurate (but less useful) to say that no element of chance is involved, the cards are fixed in order in the deck, but the players have no information about this order. All these games are sequential. Diplomacy is a game of imperfect information, because all orders are secret although the position on the board is not, and it is simultaneous, which greatly complicates matters.

4.2 Playing Perfect Information Games

4.2.1 Nim

Nim is a very simple game where players take it in turn to remove a number of stones from a number of central piles. Many different versions exist, but in this discussion the winner is the player who removes the last stone. More generally the winner is the player who moves last, but in some the player who moves last loses. It has been shown that optimal strategies exist for Nim, and similar games, and these are simple enough to be used by a reasonable player.

In games starting with a single pile, players are limited to removing a number of stones between two bounds, typically something like 1-4 or 3-7. If there are fewer stones than can legally be removed, then the next player loses, unable to make a legal move. With multiple piles, players can remove any number of stones, but only from a single pile. Nim is an example of a sequential perfect information game.

In the single pile game, basic play is no more than removing a random number of stones, but cursory analysis reveals that the first player to move can easily force a win in most circumstances, and the rest allow the other player to use the same strategy to force a win from their first move. As long as a player can leave in the pile a number of sticks that is equal to a multiple of sum of the upper and lower bounds, they can continue to do so until that multiple is zero, when they win the game. If the first player to move can do so, then they can win regardless of their opponents actions. Equally if the lower limit is a and the upper limit is b, then leaving $x + n * a + b$ stones, such that $0 < x < a$, will result in a win in n turns.

The multiple pile game strategy is slightly more complex, but still within the grasp of a human.

Nim with more than two players is more difficult. Consider three players A, B and C. At some point player A will be able to let the next player, B, win; or stop B from winning only for C to win instead. But once a player can no longer win, as the remaining players will take the rest of the stones, and there are too many stones for the player to take in this turn, there is no reason to prolong the game further. So A might as well gift the game to B. But if A and B work in partnership, they could reasonably expect to win every game. A will gift B at every opportunity, and B will play to stop C from winning. But C chances of winning have dropped from a third to zero. Now if A has agreed to work with B against C, and with C against B, A can expect to win far more than a third of games. But when B and C realise that they are both playing for A's benefit they are likely to play against A instead.

4.2.2 Chess

The minimax method is used for chess, draughts and many other games [6]. It seeks the least worst move, which will produce the smallest disadvantage if the opponent finds the perfect reply.

It models the entire game as a tree, with moves leading to new nodes, and leaves being final game ending positions. All possible moves are thus represented. Each level of the tree represents a turn, and turns are spilt between two players, here named Max and Min, where Max is the program and Min is the opponent. Two levels correspond to a 'ply', or one complete turn. If all moves from a node lead to a leaf, the node is evaluated as the maximum of its children if it is a Max turn, and the minimum if it is a Min turn, as the opponent is assumed to play the best move possible. Working back up the tree in this fashion gives scores to all the nodes. The program then picks the move leading to the node with the highest score. This program would only select moves leading to trees that end as a win, for any possible opposition reply. In practice the entire tree cannot be searched for any reasonably complex game, so leaves are created by evaluating each position on the tree at some cut-off point.

The score, or 'utility' of a position is traditionally $+1$ for a winning position, -1 for a loss, 0 for a draw, and some intermediate value for a leaf generated by a cut-off point. In practice the actual values vary depending on the evaluation function and how the value

is stored. Some games naturally produce different values, particularly multi-player games with several losers. It may be advantageous for the total utility for all players at a given time to be zero, a 'zero-sum' game is any where a gain for one player directly corresponds to a loss for the others. Almost all traditional games are zero-sum, but many real world problems that game theory is applied to are non zero-sum, and all players can gain from working together. A common example concerns the interrogation of two prisoners. If one confesses, and implicates the other, he will receive a lighter punishment. If both confess then they may or may not receive a lighter punishment (depending on the exact description of the game), but both will be punished. But if neither confess, the police do not have enough evidence against either of them, so they both go free. Diplomacy is essentially zero-sum, as the outcome depends on ownership of the 34 supply centres. But in the initial stages many of these are neutral, so a gain is not necessarily at the expense of another player, particularly as many neutrals cannot be immediately contested by an opponent.

Deciding which side has the upper hand in chess is quite easy. A basic scoring system is introduced to players at an early stage, as a rule of thumb to decide whether an exchange of materiel is should be pursued or rejected. From here the next consideration is the position of pieces. Good pawn structure, a protected king, and pieces closer to the centre of the board that are more immediately useful than those sitting on the back row, blocked in by friendly pieces. Exchanging a knight to win a rook is almost always a good idea, but a rook on the back rank, hemmed in by its own pieces, will not become important for a while; and the knight is clearly in a useful attacking position with its mere presence forcing the opponent to use another piece to defend the rook. As the scoring system is refined, play will improve.

Programs based on this method must search as many moves ahead as possible, in chess this is typically at least seven and often more than twelve. This is comparable to skilled humans, but they only consider a few paths through the tree, but the computer must consider all of them. This requires an ability to decide what the outcome is, decide which outcome is best, and the speed to search a suitable number of moves ahead. The number of possible moves for each player at each turn is the 'branching factor', b. The number of nodes to be searched is the product of the branching factors for each turn that is searched. If d is the depth of the search, then the number of nodes to be searched is

$$b^d$$

Chess typically has a branching factor of about 30, and a depth of 7-12, but most programs are designed to increase the depth at critical points. Draughts has a branching factor of about 10-15, as pieces are far more limited in their movement. Games that are proving resistant to computer programs typically have very high branching factors, as these quickly make the search unfeasibly large. The branching factor in diplomacy is the product of the number of possible moves for every piece. 34 pieces, each with 5 moves, a very conservative estimate, produces over 45 million possible combinations of moves. Allowing them 10 possible moves, slightly high, but including hold, all possible support and all possible convoys produces $b = 34^{10} = 2 \times 10^{15}$, or two million billion combinations of moves for each turn. Of course a search could be limited to including those pieces close enough to be involved against the power considering its moves.

Searches require powerful computers, but also efficient programming. As the search is of exponential size, it requires a powerful technique to reduce significantly the number of nodes to be searched. The branching factor is relatively constant throughout the game.

However there are some highly effective techniques to reduce the total size of the search. Alpha-beta pruning stops searching a tree as soon as it becomes apparent that the tree will not be useful, rather than searching the entire tree to a conclusion. It reduces the number of nodes searched to approximately the square root of the total, allowing the search to be twice as deep in a given time. Progressively deepening the search finds the best move can be found in a strict time limit. This time the exponential factor is an advantage, because a search to depth $d-1$ is far faster than the subsequent search to d ply. Now further gains are made by by considering trees that led to the best result in the last search first. The alpha-beta pruning can then prune the alternatives faster. In addition trees that lead to the same result can be combined. The critical trees can then be selectively searched by some extra ply, to avoid a disaster appearing but without having to search the entire tree as deeply. The key figure is the rate at which moves are searched. Being able to see further ahead than an opponent is a crucial advantage, as it may only take one less than perfect move to lose an expert game. Sheer speed is not a huge consideration if playing diplomacy by email, as moves can be decided over hours or even days, but it would be inconvenient for the scope of this project if testing took days for a single game.

However chess programs are not merely a good search algorithm. Libraries of openings are available, roughly equivalent to what a Grandmaster memorises. Thus neither human nor computer has to waste time in the early stages of a game when move and counter move are already known.

Databases of endgames are also available, but less common. The first chess machine[7] challenged players to draw with a king, against king and rook, and could win from any position. Other algorithms are available for other endgames. However most top level games of chess end before the end game, and the search algorithm performs adequately enough that these are less common in chess programs.

Similar theory is available in Diplomacy, but few openings stretch beyond Fall 1902. As seven players are involved the combinations of plausible moves very quickly become too big to cover. End game theory is potentially more useful. Lists of stalemate positions are available, and enable countries to force draws if correctly used.

4.3 Imperfect Information Games

4.3.1 Games of Chance

Backgammon requires the roll of dice to determine what moves are possible for a player. This makes the minimax search slightly more difficult, as only a subset of the nodes searched will be possible moves after the dice have been rolled. However it is easy to add an extra layer, representing the probability of a roll. The utility values become expected values, and the principles of the minimax search still apply. This is known as an expectimax method. But perfect play can be defeated by better luck. A backgammon program has beaten the world champion, but the programmer acknowledged some fortuitous dice rolls.

Almost any game of dice or cards can be treated in this fashion, though like perfect information games, the game tree may be too large to search adequately.

4.3.2 Games without Chance

These can still be complex, even with very simple rules and no bad luck to guard against. Consider the children's game Paper, Scissors, Stone. Here both players have three choices, and each choice will win against one, lose against another, and draw against itself. The game lasts a single turn, so the search is only a single level deep.

Now consider searching for the best move. Assume all opponents moves are equally probable. Each move now has a one-third chance of a win, a loss or a draw, but all are equal. Any search fails, because no move is identified as better or worse. But crucially cards and dice have no memory, but a human opponent does. All opposing moves are no longer equally probable because a human is generally reluctant to pick the choice that lost the last game, and do not want to pick the same move twice. The probability of choosing each move is not equal. But any analysis will show that it is purely a game of luck, and a computer choosing randomly will win, lose and draw one-third of the games that it plays.

But it does have useful real world applications. A football player steps up to take a penalty kick when he sees a former team-mate telling the goalkeeper where he normally aims his penalties. Should he change his mind? Is the information as useful when everyone knows the goalkeeper has it? Now the player is ordered to retake the penalty kick, should he aim the same way again, or change his mind? Should the goalkeeper dive the same way again? If the goalkeeper has dived the same way for the first three kicks in a penalty shoot-out, is he more or less likely to dive the same way for a fourth time? If a companies share price rose yesterday, is it now more likely to rise again today?

These problems can all be stated in terms of probabilities, but the probabilities cannot be explicitly calculated. Statistical analysis of all previous similar situations will help, but might not be available.

In Diplomacy there are many situations that involve two players guessing how the other will play. For example consider a Turkish fleet with support attacking an Italian Fleet in the Ionian Sea in a Spring turn. The Italian fleet is dislodged and can retreat to Tunis, the Tyrrhenian Sea, Naples or Apulia. The Turkish fleet can attack any of these provinces in the fall. Assuming Italy controls Tunis, it has to defend Tunis and Naples, so without any other nearby units Italy must retreat to the Tyrrhenian Sea, from where it can bounce in either Tunis or Naples. If it guesses wrong then Italy loses a centre, and will have to disband a unit. If Naples is lost then Rome is threatened as well, and Italy could lose two centres. Clearly Italy will favour going to Naples. But this is equally apparent to the Turkish player, who can move to Tunis, and still take a centre. So Turkey clearly gains more often from the move to Tunis, so Italy can expect a move to Tunis, and counter it. If both move the same way, they bounce and the problem occurs again next turn. Should Turkey attack the same centre again, or attack the other. If no other units are close enough to interfere, this could last for several turns. Turkey's optimum move is often to the Tyrrhenian Sea, as Italy will move out of it to defend one centre, and can do nothing to protect the other in the next turn. But it is a brave Italy who holds, and a highly embarrassed Italy who holds when Turkey has moved straight for a centre. The fleets could have even more options. If Turkey can support another unit into a better position (or convoy an army into Apulia), while Italy has tried to defend a centre, then Turkey has gained a small advantage, and the Italian fleet can no longer defend both centres so Turkey has a bigger advantage to come.

4.3.3 What Price has Information?

Information inherently has utility. For example, in bridge, players bid to play the contract, and the bids help describe players' hands to their partners, but also convey information to opponents. One half of the partnership, usually with the better, or more unusual hand, benefits more from keeping quiet and disguising their hand, and then choosing the contract,than the other, as the hand belonging to dummy is visible to all.

Once play begins declarer will wish to hide his hand, and so plays cards in an ambiguous fashion whenever possible. So the highest of touching cards should always be played, as this disguises the position of the others. If the lower card is played, and wins, it suggests that the opponents do not have a higher card. Equally the defense should play low, as this signals to their partner that they have the higher cards. For instance, as the defence, playing the queen, from king and queen, suggests to your partner that you have the king when declarer wins with the ace. Equally if declarer is holding the ace and king, then it is better to play the ace than the king, because both will win but rising with the ace keeps the location of the king disguised. But at times it is better to ignore these rules of thumb to sow confusion in the opponents.

Equally, in Diplomacy, many opening moves are promoted because they are ambiguous, and can be represented as all things to all players. In Spring 1901 Russia and Turkey often bounce in the Black Sea, Russia and Austria in Galicia, Austria and Italy in Trieste and Venice, France and Germany in Burgundy, or occasionally England and France in the English Channel. But none of these are a necessarily a declaration of intent against the other. Often they are prearranged, as the unit has nothing better to do, and it is safer to bounce than risk the failure of a demilitarised zone. The English Channel, along with the likes of Tyrolia, Rumania, Piedmont, Prussia and Silesia are less favourable locations for prearranged bounces, because the units involved could all be more profitably used somewhere else. But the advantage of the arranged bounce is that the players concerned can represent it as an attack (by either player) to potential allies, and disguise their true intentions for another turn.

Poker [8] often sees players with a potentially good hand risk money to find out if they will catch the card that gives them the winning hand. Texas Hold'em is regarded as more skill than luck compared to other poker games. Players are each dealt a pair of cards then bet. Those who are still in see five common cards dealt face up in the middle of the table with betting rounds occurring after the third (flop), fourth (turn) and final (river) cards.

The four betting rounds before players know what their final hand is can be expensive, but are a perfect example of putting a price on information. Players with good cards (a pair, or two face cards) at the start of a hand bet large amounts before seeing the flop to discourage other players with worse cards from staying in and possibly winning. Players with potential hands, like two cards that are touching or suited or both, want to see the flop, as the three cards will tell them if they have a chance of winning the hand. Players with low, unsuited cards want to get out of the hand as cheaply as possible, because their expected return is so low.

A player with an open-ended straight draw (four consecutive cards) after the flop has eight cards in the deck that would complete the hand, but a player with an inside draw (any other four from five consecutive cards) has only four. Neither player expects to win, as their hand is practically worthless unless they can catch the missing card. But if they do catch it, then the straight is high enough to make losing unlikely. Now if enough opponents are involved then the total pot might be large enough to make play worth

while. Here there is a definite price to be paid for the information, but if the price is lower than the probability of catching the missing card multiplied by the difference in the cost of folding now compared to the gain of taking the total pot, the price is worth paying. Because Texas Hold'em features open cards, the potential of the opponents hands can also be gauged. The straight, given as the example earlier, looks even better if the common cards have no pairs, or three or more of the same suit, because a straight is the highest poker hand a player can make without these cards.

But the stereotypical poker play is the bluff: betting more than the opponents can afford to call but holding cards that are unlikely to win. Very big bets accompany weak hands, strong hands bet smaller in the hope of attracting a call, and gaining more money. A player with a reputation for bluffing risks losing big because opponents are more likely to call, but a player with a reputation for playing straight cannot steal pots as often. Other players prefer to trap, playing a big hand as a far weaker one, and encouraging opponents to bet into them before raising big to take an even larger pot. The solution lies in knowing your opponents, who bluffs, who calls the bluffers, who traps, and who gets suddenly excited with a big hand.

Again in diplomacy you must know your opponents. Will they hold to an agreement or have they already made an alliance someone else? This is more difficult in no-press games because the communication between players is restricted. The problem is that the number of significant events per game is too low to allow a statistical picture of the opponents to be built. Poker players can play hundreds of hands over an evening, and see a bluff occur every few hands, but the diplomacy program will see only a handful of stabs in a game.

It is almost impossible for the program to tell if the stab was because the ally can never be trusted, or because the ally had always intended to attack Italy next and it is unfortunate that the program happened to be playing Italy, or just because the program left to good an opportunity to be missed.

4.4 Approaches for Diplomacy

Three distinct approaches were identified to find a suitable set of moves that the computer could play. The first was a brute force exhaustive search of the move space, examining all possible combinations of moves by all players (chess program style), and picking the best. The second was to look at each unit in turn, and pick the best move for the unit (probably closest to human style play) and the third was to identify where the computer should want to move, then work out how to get there with a location based approach.

4.4.1 Exhaustive Search

The brute force approach relies on being able to search all possible moves (ideally over several turns) and identify the result of each. The best move is the one with the least bad possible outcome, as this approach assumes perfect play by opponents. Each node of the game-tree is assigned a score, with the path through the tree determined by both sides maximising their own scores. Essentially any search trees possibly leading to defeat are avoided, this only leaves search trees that lead to wins, or draw against perfect play. Of course if the search is not deep enough there could be a nasty surprise at the end.

Finding the outcome of a set of moves in diplomacy is quite complex, as all moves are potentially interlinked and are simultaneous. However it was solved by the judge

software, and the necessary algorithm is freely available.

In Diplomacy the simplest measure of success is obviously the number of supply centres controlled. Other considerations could be the number, and strength, of your apparent enemies. A power with sixteen centres is in an excellent position to win, but if all the remaining opponents have united against it, it might need to be lucky to draw. Equally important is that taking a centre of an existing opponent could be more appealing than taking two from someone else. But equally if an ally is occupied far away then it could be the perfect time to stab, as the ally can do nothing in reply. Taking an early lead causes jealousy, and strong starters are liable to draw attention to themselves and get attacked from all sides. A suitable score could be the difference between the number of centres that you control against a weighted average of everyone else, with the weightings representing the degree of friendliness towards your rivals.

What this method does add is long term planning. The program can reach moves with no immediate advantage, but which will prove useful a couple of turns later. This is obviously useful as units act towards a goal some distance in the future, but becomes increasingly important in the later stages of the game when a player is building units in home centres then having to wait several turns for them to reach the front line.

Because the future potential of any move now is always considered by the search, it means that the scoring of a position can be purely short term. Ultimately if a winning position has a score of $+1$ and a losing position of -1, and the search is infinitely deep (reaching either a game finishing or a previously encountered position at the end of all branches of the game tree) then all other scoring can be neglected and the search will discover it for itself.

Also the program simply has to know the rules and have a good evaluation function, no further information needs to be provided. The program will deduce tactics for itself. Both self-bouncing and beleaguered garrisons would be found when appropriate. Additionally it would ignore moves that will never produce a benefit, like self-bouncing when there is no threat. It will also realise that a location bordered by a single enemy unit is only threatened if the occupying unit away, not if it moves towards the enemy. These cases are all very difficult to otherwise provide algorithms for that only work when appropriate.

It is clearly a very powerful method, but suffers from some weaknesses. One is that Diplomacy is not a game of perfect information. Finding the best move requires knowledge of what your opponents' intentions are. Because the search assumes that an opponent will play perfectly, and so looks at the worst result of a set of moves it may never take the risks that are required to advance. The second is that it is a game involving seven opponents, and some may be happy to help you. This means that a stronger position is not stronger if it antagonises an ally.

The method lacks randomness, it exists to find the theoretical best move, so everything else must be inferior. If this best move can be guessed by the opponent, and countered, it is clearly worse than an unexpected move that still improves the overall position. But playing a less than perfect move against its perfect defence clearly cannot be any better either. This is the weakness of the search in a game of imperfect information.

Consider a situation where a player has a single unit trying to defend two centres from a single opposing unit. The player succeeds if he moves to the same province and bounces, but fails if the enemy moves to the other province, taking the centre. This is essentially a 50-50 chance of keeping both or losing a centre. However the search will always assume that the centre will be lost, as against perfect play the guess will always be wrong. So defence of the centre will lead to its loss as surely as ignoring the threat, and

doing something completely different would. Of course the outcomes could be weighted by the probability of the opponent making the perfect move, but this leads to the problem of estimating the required probabilities.

Equally a 50-50 guess to win a centre will have a similar result, as against perfect play the centre cannot be won. Actually the centre is doomed, as the defence will guess wrong eventually if it cannot change the situation.

Also the search assumes that everything that improves your position is good, and only considers the final position of the search, as the method used to get there is unimportant. If this upsets an ally then it weakens the actual position. For example, a French fleet in the English Channel, or a German fleet in the North Sea are both nice, occupying useful strategic positions guarding a flank, and threatening or supporting the low countries. But both are also hugely disturbing to England, who will not remain an ally for long if confronted by such a potential threat.

The chess style search is naturally pessimistic, but when every country has the threat of being attacked by an alliance of two or three neighbours, with little prospect of survival unless they can break up the alliance, it seems that the best moves, to protect against the threat of overwhelming attack, are those that are either unduly defensive, or threateningly preemptive.

Ultimately the search method fails because Diplomacy is a game of imperfect information. A search can be profitable in games of chance, as probabilities can be incorporated into the score of an outcome, to give an expected value. But an expected value for a set of Diplomacy moves relies on having a good estimation of the probabilities of each of an opponent's moves, and this is difficult, as these in turn depend on, and are modified by an opponents estimation of your own move.

4.4.2 Unit by Unit

The natural human approach is probably to split the units up into groups that are close enough to interact. The player will have long and short term objectives for each group and go through the possible orders, and possible counters by the opponents to find some suitable set of moves.

While it is simple to consider each unit in turn, and find a move one at a time, it is difficult see how a concerted attack could be made within the confines of such an algorithm.

4.4.3 Location, Location, Location

This method aims to consider the current state of play and provide a set of moves aimed at improving the immediate situation. The long term plan is largely ignored. It scores every location on the map, then units can move to the highest scoring provinces.

Scoring individual locations is easier than entire positions. Among the factors to be considered are whether it is a supply centre (more important in fall than spring), who controls it, whether it is threatened with attack or whether it can be defended by an opponent and how important the neighbouring locations are. A small random element added to each location score introduces some unexpectedness. The maximum size of random element compared with the difference between the locations scores allows a precise probability to be worked out, so that one of two identical locations is favoured exactly 50% of the time, but a small difference in scores will lead to a 60-40 or higher

ratio. It also means that some moves are still obvious. An alternative would be to move according to the ratio of all the possible scores, so one with a score twice as high is moved to twice as often. This would mean all moves with a non-zero score have non-zero probability. For example, in Spring 1901 England generally moves her fleets to the North Sea and either the Channel or the Norwegian Sea, the former threatening France and the latter Russia. If this was entirely predictable then France and Russia could both play to counter, knowing that it was about to happen. So the two locations must have similar scores so that England picks the Channel over the Norwegian Sea somewhere between 70–30% of the time, although the exact ratio would be determined by the scores of the two areas, and the size of the random element. But it should never play the army to Wales without the fleet moving from London to the Channel.

The location score is currently 0 if the province is unthreatened, and 25 if an enemy threatens a supply centre. If an enemy supply centre is threatened it is worth 25, or 40 if it is undefended. But an enemy home centre is only 20 if threatened or 60 if it is undefended, because although it is inherently more valuable, the enemy is more likely to defend it if they possibly can. This is the Naples or Tunis dilemma discussed earlier. To this another 5 is added for each enemy adjacent to the province, or 10 if they occupy it. This is to encourage attacks against enemy units, cutting support or capturing their province. To encourage units far from the enemy to move towards them, each province between them and the enemy is worth -3. And each province also receives a bonus of a third or a fifth of the adjacent provinces scores, to reflect the potential for next turn, depending on the season. Finally a bonus of up to 20% is added, to ensure a random element to movement if two scores are sufficiently close together.

The next step is to give each unit a list of the locations it could move to, sorted in order of their score, and including the current location so that they can hold if they are already in the right place. Each unit attempts to move to the highest location on their list. Before the orders are finalised there is then an algorithm to resolve conflicting moves, so that if two units are in the very common situation that they are both attempting to move to the same place, one will either change to move to the best alternative or support the move instead.

There is now the possibility of using fixed or variable location scores. It would be possible to work out a score for each location, or even a score for each location for each power, and hard code how important the location is, or load it in from a file when the program starts. Alternatively a formula could be used to calculate the score for each location. In either case the basic score can be further modified depending on the location of enemy units etc.

Variable scores were chosen because it was more flexible to include a formula and calculate scores on the fly. For example, Burgundy is important to France at the start, but useless once French forces have moved beyond Munich. This was easier to represent by a formula that treated it simply as safe or threatened, and included the distance to the nearest enemy to encourage units towards the front lines.

Essentially this algorithm decides where best to go, then how to get there by matching the current situation to a list of particular cases. Its obvious weakness is that it is very short term. The potential of future moves has to be included in the score for a location - mainly by including the scores of neighbouring locations in the formula, and units can get stuck far from the action if they have no immediate role. It is also purely tactical, designed to take locations it currently identifies as important. This means that it also has to incorporate strategy in the scoring of locations, to avoid upsetting friendly powers.

4.5 Strategy

The methods above handle the tactics, questions like how to use three units against two to best ensure occupying a location, and work in the two player game. They do not consider when alliances should be made or broken.

The basic strategy is to attack the strongest power, as it is most likely to win. If this was the only strategy for every power then the game would never end. But nobody would lose either.

The next consideration is to attack the weakest power, and remove them from the game. This improves the status of the draw if no power can win out right. Additionally it is better to pick on the weak because they are least able to fight back. A power that is heavily outnumbered has little chance of taking centres in a counter attack, because the attacker has the resources to cover all his centres. If the weak power is about to be eliminated, then grab a share of the spoils rather than leave it all to others. Finally, once the weaker power is eliminated, they can never fight back or stab, and the other powers have little claim to the home centres of the eliminated power.

If two weak powers are in the game, and one is crushed and eliminated, then the other follows, this is better then two very weak powers remaining, both trying to regain their home centres, and open to any alliance that will help them stay in the game. Equally it is more efficient to fight on a single front, than try and fight two weak powers at the same time.

If the emphasis was on fighting the weak before the strong then there is a risk that another power wins while you are fighting a weak power, but making slower gains. However there is an argument for ignoring the strongest power until they reach a certain threshold, at which the remaining powers should unite against them. But at the same time is is common for any power taking an early lead to merely encourage the others to unite against them, although they may only have six or seven centres. If Germany or France can control 6 centres in 1901, or 7 in 1902, then the other will often ally with England. So a slower start, splitting the centres five each, with Belgium being given to England to avoid a more heated dispute, is relatively common, and often better in the long term than a swift grab.

The current algorithm starts by halving all current scores, so that past transgressions are forgotten. It then adds a distance score to each, equal to 10 minus the shortest distance between the two powers. Enemies with no units count as having a distance of 100. This encourages it to move against a nearby power.

Next step is to find number of powers with the largest and smallest number of units. A score of 30 is split between the largest powers, and 20 between the weakest powers. This helps stop the strong but eliminate the weak.

It then checks all home centres, and if any are occupied by an enemy then a bonus of 60 is applied to their rating, so that powers will almost always fight to regain their home centres.

Other events will also modify ratings, so they are increased by 10 if a unit is dislodged by the enemy, another 10 if they capture a province, and 10 if they lose a province, so that counter attacks are launched, but if an attack is successful, the momentum is not lost.

The enemy with the highest score that is nearby is attacked first, but if the enemy is outnumbered by the attacking nation, another enemy will be attacked as well, to help eliminate small powers, without tying up large powers against a single small enemy.

However it still allows stabs if an ally suddenly becomes very strong or weak, or leaves a border weakly defended.

More complicated criteria will give better performance, perhaps a power is content to remain in second place for a while, as long as they are not to far behind. So a 7 centre power will not attack a 9 centre power, preferring to eliminate the weak, but a 14 centre power will attack a 15 centre power, as the enemy is closer to winning.

However there will always be exceptions, and it is not clear that the strategy algorithm performance will match an increase in complication.

Chapter 5

The Program

5.1 Requirements

The program was written in C++. The aim was to be able to play the full game of Diplomacy, at a novice level. The other requirement was for the input and output to be compatible with the judge software for it to be able to play in email games with a minimum amount of human intervention.

The definition of minimal intervention is loose. It was intended that a user would cut and paste judge results from emails into the input files, then the appropriate orders back into an email to be sent to the judge. It would be possible for the program to communicate directly with the judge, but the additional complexities of providing the program with an email address and email handling code were not deemed within the scope of this project. It would also make it more difficult to hand adjudicate test games, and the two and five player test games would require variant files to be written for the judge software.

The decision was made to use the location based scoring, and calculate the location scores on the fly. The search tree was too large, the extra code to calculate outcomes for the search would be considerable, and there was no method to calculate even approximate probabilities for enemy moves to base the expectimax search on. A location based approach could easily be prototyped, starting from random moves and gradually improving the location scoring and order-making code.

5.2 Architecture

The program classes correspond closely to the board game. Each province is an instance of a *Location* class, holding information such as name, abbreviation, owner, occupying unit, distance to other locations, and pointers to bordering locations. The map is a *ListLoc* class, holding an array of locations, and methods to parse the map file to create the locations, and search the array.

Each player is an instance of *Nation* which includes a list of *Unit*s. The *Nation*s are grouped into *NationList* which has methods to run the phases of each turn. There are other functions concerned with parsing the input files. A *order* class separates the actual orders from the units, but this is a legacy of the program parsing the judge output to recreate the orders from the last turn. As every unit has a single order, this is an artificial distinction and *order* could be incorporated into *Unit* without loss.

The units are all destroyed before the spring and fall movement phases, and then recreated from the input file. Retreating units are then destroyed and recreated from the input (unless they are disbanded then). The build phase sees new units created, and disbanded units destroyed before the next movement orders are created.

5.3 Development

Development consisted of producing prototypes, and gradually increasing the functionality offered. The first was to parse a basic judge format output file and extract the nationality, type and position of units, along with details of their last order. Next the program was modified to parse a judge standard map file, with gives details of provinces, including name, abbreviation, type and bordering provinces. The unit details could then be matched to the board.

Then methods were added to produce retreat orders for any units that had been dislodged. This required finding scores for all the border provinces that were valid retreat options, and picking the best. This was extended to produce moves for the next turn, so that the program could produce a move order for all units.

The next step was to give units a list of target provinces for the move, and compare all the units of the same nation to avoid several units moving to one province. In the event of two units moving to the same place, one will change to either move elsewhere or hold. Any units still holding after the first pass are then checked to find if they can support another move, or should move elsewhere.

The retreat methods were rewritten in a similar style, to avoid a power retreating two units to the same location, as both would be automatically disbanded.

Another group of methods are used for adjustment phases, to change ownership of supply centres and to build or disband units as required. A save/load option can store and recall the ownership of all the supply centres, which combined with the last results file is enough to recreate any point of a game.

A simple text based menu allows the user to play through an entire game, inputting the judge files and receiving the next set of orders for all sides.

At points the input parsing was extended to cover bicoastal provinces and different styles of adjustment orders while explicitly ignoring the extraneous information included in the judge report, such as the judge details, the next deadline, and the lists of supply centres and dislodged units which are both determined by the program from the main results.

The location score and movement producing methods were evolved as testing revealed situations that had not previously been considered.

There was a small set of constraints applied to the location scores. These were essentially that an unthreatened centre has a score of zero, that a home centre is more important to defend than other centres, and that a home centre is correspondingly more likely to be defended. A lightly defended centre is more attractive than a heavily defended one, and so undefended centres are most valuable to the attacker, and home centres least. Other bonuses were applied if more enemy units were near a centre, to encourage units to form centres of mass large enough to make gains. Bonuses were applied depending on the closeness of the centre to the enemy. Finally a portion of the neighbouring location scores were added to represent future moves, and a random element to produce unpredictability.

The program was tested using the two player game described below. The algorithm to support another unit was expanded to cope with more situations including: where one

unit is holding and another is moving into it; where three or more units are moving to the same place to ensure that the one with the worst current score moves, maximising the increase in score; where a units move depends on another unit successfully moving out of the target location to avoid bouncing and when a unit already has as much support as it requires to guarantee the success of its move.

The next stage was to add an overall strategy for each power. In the two player game it was sufficient for the program to maximise each power's advantage, but in the five player game it was necessary to concentrate on some enemies and make peace with the others.

Each power maintains a score for all of its rivals. The scores are increased if the rival is the strongest or weakest remaining in the game, and if the rival occupies a home centre, or has just occupied a province that the power controlled. Scores are also modified for the distance that each rival is from the power. The score is also increased if the power has just taken a centre from the rival, so that the power can press home a successful attack.

The power then targets the rivals with the highest score. If it outnumbers its highest rival, it will also attack the next on the list. If it is too far away from a rival to attack, it will also attack the next on the list, so it will always have a rival to attack. The scores are halved at the start of each turn, so that past transgressions become less important, but will not be forgotten until another rival becomes more important.

If a rival is not to be attacked, the scores for its provinces, and those containing its units are held at zero, so units will not attack 'friendly' powers. However this does not apply if the friendly power is outnumbered. So a friendly power can still be stabbed if the circumstances are right, and it might suddenly top the list of rivals if it becomes the strongest or weakest power.

The next step would be to decrease scores on receipt of messages of support, allowing powers to form alliances and even coordinate plans instead of simply not attacking.

Because the orders are not adjudicated internally, the program recognises only that units bounce, or are dislodged, but not which unit bounced with them, or supported the attack. This makes it difficult for the program to update the strategy against an opponent who is bouncing or supporting attacks against them. The strategy is updated when units are dislodged, as they know who dislodged them (and cannot retreat into the province from which they have been dislodged). It is also updated when locations change ownership, as again the program knows explicitly which powers were involved.

5.4 Testing

The parsing program was tested against a number of judge results files until all orders were handled without errors. The judge results files were from the Vermont Group No Press Tournaments, and these are available from http://doug.obscurestuff.com/vermont.html

The two player and five player versions of the game were used to play test games, and errors corrected as they were found. The test game results were also used to optimise the constants in program algorithms, and used to decide which additional features to add. For example, the algorithm for finding chains of moves was added after a test revealed that units were consistently bouncing because another move had also bounced.

Chapter 6

Two Player Game

6.1 Two Player Variant

The program was initially tested for a two player game using a simple map with no sea areas. The south-east corner of the board was used with Austria and Russia as the opposing powers. The north and west edges is the line including Moscow, Livonia, Warsaw, Galicia, Vienna, Trieste, Albania and Greece. The Aegean Sea and Black Sea become plains, and the Eastern Mediterranean and Syria are excluded. To balance the distribution of neutral supply centres, Serbia, Constantinople and Ankara are no longer supply centres. The fleets starting in Trieste and Sevastopol are replaced with armies.

Figure 6.1: The two player map

The map is biased slightly towards Austria and, in the absence of luck or a third power to ally with, the game will be won by the first power to gain an advantage in the

number of supply centres. Austria can move Trieste to Serbia and use Budapest to deny Rumania to Russia in Spring 1901, and then choose to attack Rumania again but now with support in the Fall, expecting to bounce and so nothing changes, or attack Bulgaria (or Greece, but that is far weaker play) from Serbia so that both gain a build.

It is stronger for Austria to delay taking Bulgaria until a Spring turn. In the Fall Russia must support the defence of Rumania against the possibility of Austria attack. But it gives Austria the chance to support a move into Galicia instead. From there additional pressure can be brought against Rumania. However, because an successful attack on Russia is not required, the long term advantages of the move will not become apparent.

Unfortunately for Russia, there is no way to gain the centre advantage against Austria. The Austrian army built in Trieste will move to Serbia. Here it has the flexibility of threatening Rumania but also able to take Greece. The Russian build in Sevastopol can reach the final neutral in Smyrna by the next Fall by moving through Armenia, but this allows Austria to take Rumania, gain an army advantage, and still take Greece later, a position Russia cannot recover from. Alternatively it can move to the Black Plain. Here it can protect Rumania and threaten Bulgaria. But if Austria moves Budapest to Rumania, cutting support, Serbia to Bulgaria, bouncing with Black Sea, and Bulgaria into Greece, at best Russia can only hope to hold Rumania or take Bulgaria. With the Bulgarian army retreating into Greece, Austria again gains a centre advantage.

Galicia is an important province for helping attacks into Budapest or Rumania. The armies in Warsaw and Vienna are doomed to bounce endlessly here unless someone can find a spare army to support them with. Having occupied the province any support given from it can be cut so it is better to use the army in it to try cut support yourself at the risk of losing the province. However an enemy attempt to take the province will require support drawn from somewhere else (Ukraine or Budapest), and once they have taken it they have the same problem of using it only at the risk of losing it. Taking Galicia appears to be only at the cost of a better move. But Vienna and Warsaw have no better move than to deny Galicia to the other.

6.2 Two Player Results

The first tests of the program with the program playing both sides did not allow support orders and did not feature the random element of the location scores. These finished as a draw because neither side could dislodge the other's armies but showed the locations were being rated properly by both sides. The next test games with support implemented show quite sensible play, but a defensive tendency. This was because units were getting a higher score for defending supply centres than attacking enemy supply centres. If two adjacent centres are occupied by different powers then they will defend (hold) even though they cannot possibly lose the centre by attacking the only adjacent enemy unit. Multiple adjacent centres would see several units mutually supporting each other.

A more aggressive location scoring method, with locations occupied by enemy units getting a higher score, lead to more aggressive play. Most tests led to Austrian wins, as expected, but with varying degrees of competence. Some tests resulted in repeated positions, as all moves by both players bounced or saw units moving back and forward into a province. In particular Russia would build in Moscow, move to Livonia, then back into Moscow repeatedly. This was eliminated when the random element was added to each score, so eventually the cycle should break.

The random scoring is most noticeable in Fall 1901 where, as described above, Austria can support an attack on Rumania, so all units bounce, or split its forces between Rumania and Bulgaria, so each gain a supply centre. In tests Austria has favoured initially attacking Rumania, often repeating the attack in 1902 before changing to take Bulgaria instead.

Another common tactic was for Russia, having finally taken Rumania and built a new army in Sevastopol, to use Ukraine to support the new army from Sevastopol into Rumania, while the army in Rumania attacked somewhere else (normally Budapest). This is an example of the "Beleaguered Garrison" tactic that emerges from the scores generated. It ensures that Rumania is defended regardless of the result of the attack by the army already in Rumania. Unfortunately for Russia, Rumania is too heavily outnumbered and falls to an attack with two support. However it would be an elegant way to stop an Austrian supported attack on Galicia, as it would cut support from Budapest, or, if the army in Budapest moved, it would walk into a undefended centre. Equally if Rumania attacked Serbia, it would often trap the newly built Austrian army in Trieste. But Austria can always support a move into Serbia using the armies in Budapest or Bulgaria.

In a test of Austria against a human opponent, Austria lost by 1906 after Russia occupied Galicia in Spring 1903 and used it to cut all support that Budapest tried to give. As Galicia is not a supply centre, Austria never attempted to attack, always preferring to attack Rumania. As a result the army in Galicia effectively nullified the Austrian armies in Budapest and Vienna. Vienna always supported Budapest against the threat of an attack from Rumania. This was perfectly justifiable, and would have maintained the status quo if Russia attacked Budapest. But it was overly defensive and allowed Russia to swing round through the Black Plain and take Bulgaria. The resulting Russian build was then able to take Smyrna as Russia's final centre.

Chapter 7

Five Player Game

7.1 Five Player Theory

7.1.1 The Map

The five player game uses the same land only style of map, but with five players to allow the program to form alliances. The map includes all of mainland Europe with the exception of Iberia, Turkey and Greece. The Adriatic, Baltic, Gulf of Bothnia and Skagerrak have been filled in. The supply centre positions have not been changed, but Warsaw and all of Austria is neutral, and Austria (now more accurately "The Confederation of Balkan States") occupies Serbia, Rumania and Bulgaria. With 24 supply centres left on the map, 13 are needed to win.

Testing revealed the map to be quite defence oriented, with little room for strategic manoeuvre. Because the sea areas that have been removed are not supply centres, the ratio of armies to provinces is higher than the normal game. Adding the Black Plain to the map would provide more movement in the South-East, currently Russia is practically impregnable to attack between Warsaw and Sevastopol. Equally adding the Gulf of Lyon and Tyrrhenian Sea would provide more scope for Franco-Italian wars. Switzerland should also be made passable, as the Franco-German front is quite narrow, and would become a three-way battle rather than two two-way confrontations. But ultimately fleets provide the necessary flanking capability.

7.1.2 The Confederation of Balkan States

The Balkans (referred to as Austria in the test output) can take Budapest at will, and force Italy to guess to take Trieste or Vienna. In Spring 1901 Serbia will go to Trieste to expecting to meet Italy, Bulgaria can go to Rumania, expecting to bounce with the Russians coming from Sevastopol, and this leaves Rumania to take Budapest, and support Serbia into Trieste at the second attempt. But alternatively it could go to Ukraine, forcing Russia to guess between Moscow and Warsaw, and supporting Warsaw back into Moscow if it suspected Russia would self bounce. It would also stop dead an early Russian attack with Moscow moving to Ukraine to support Sevastopol into Rumania.

If Italy moves Venice to Tyrolia rather than Trieste and Rome to Venice in Spring 1901 , then the Balkan army in Budapest can happily support Bulgaria into Rumania. If Italy takes Trieste then the Balkans can retreat into Vienna. This strengthens the defence and still guarantees two builds. Alternatively Budapest can move to Vienna, or support

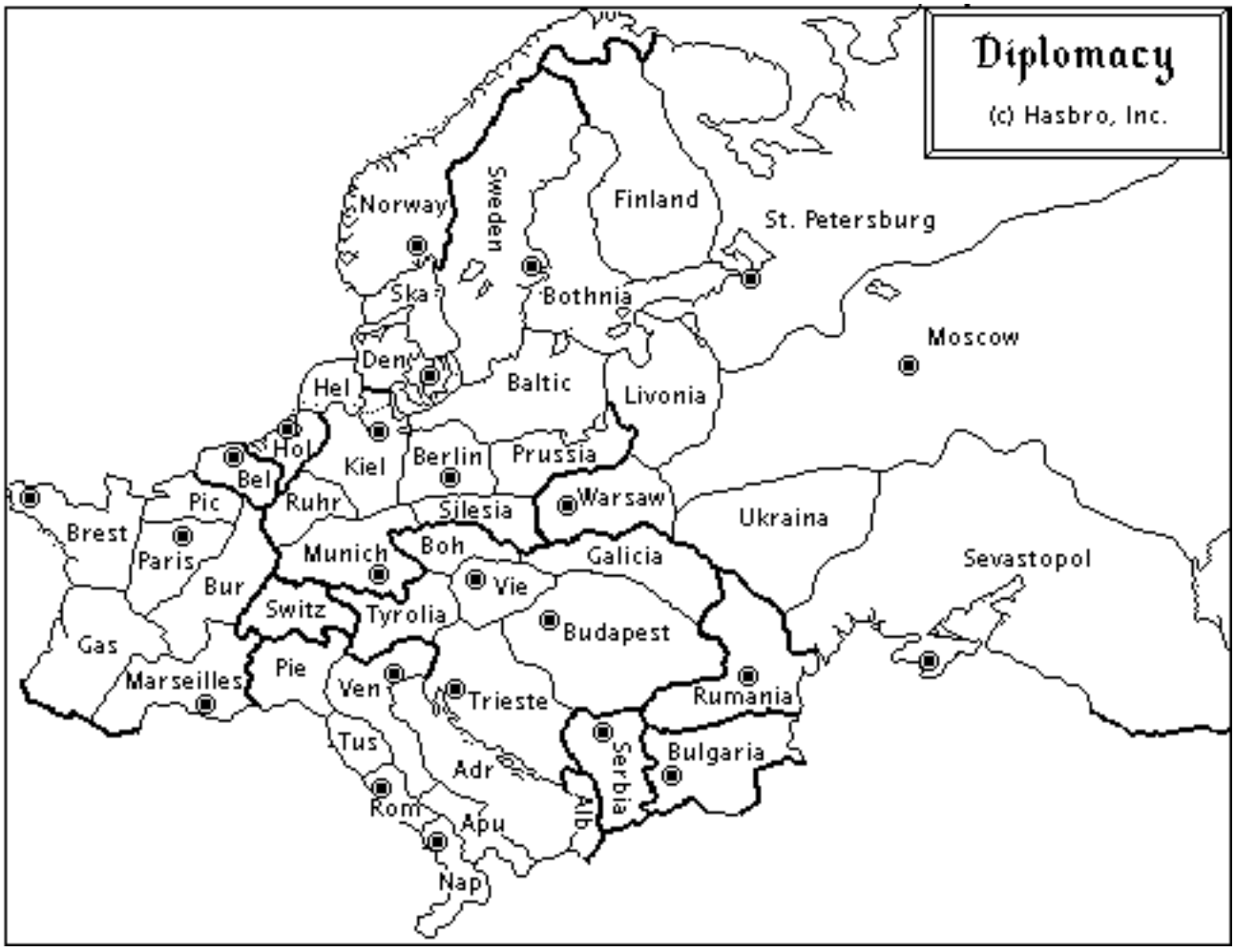

Figure 7.1: The five player map

Trieste. Again this guarantees two build, but also forces Italy to guess right between Trieste and Vienna to get a build.

To win, they will need three home centres, three Austrian and Italian, and another four from Warsaw, Russia and Germany.

7.1.3 France

France has little choice. It can attack Italy through a single province front south of Switzerland, or it can attack Germany. Belgium is the only build within reach, and its capture is quite doubtful. Spring 1901 will see Brest move to Picardy, and Paris to Burgundy, possibly supported by Marseilles. The fall will then see the attack on Belgium. If France fails here then it will never get much better.

Defensively France is impregnable if it can secure Belgium and the Ruhr. With armies in Belgium and Burgundy supporting one in Ruhr, and one defending Marseilles, no enemy can bring enough units to bear to breach the line.

Of course Germany cannot crush France without seeing Russian pincers racing through Scandinavia and Warsaw. Equally Italy has enough problems to the East that a single French army could easily mount a surprise break out into Piedmont, then an undefended Tuscany or Venice, and fall upon the rear of the Italians. Once the first is through then

others will follow more easily, and Italy has to start disbanding the units it requires in defence.

7.1.4 Germany

With the collections of neutrals in Scandinavia and the Low Countries, Germany looks strong, with a choice of which two builds to guarantee by Fall 1901. But it is also vulnerable to being crushed between Russia and France. The winning thirteen centres are three home, three Scandinavian, Warsaw, St Petersburg, Moscow and Holland is ten, and another three from Belgium and France, Sevastopol, and Vienna. The problem is that France can defend the west indefinitely if they can take Belgium and the Ruhr.

In Spring 1901 Germany must decide North or West for the army in Kiel. If North then it can bounce Russia in Sweden in the fall, a useful delay. If West then it can bounce in, or even take, Belgium, crippling France. Regardless, Berlin can move into Kiel after it, and then go the other way. Marching Berlin east will work only if Russia has ignored Warsaw, and even then gains little and is vulnerable to counterattack. Berlin to the Baltic Plain, combined with Kiel to Denmark, will win Sweden, but allows France Belgium.

Munich to Ruhr threatens Belgium, but is vulnerable to a French move to Burgundy. Munich to Tyrolia is equally vulnerable if it succeeds, but has a good chance at Vienna, Trieste or Venice. If it does fail then it has only weakened Italy, and strengthening the Balkans is not really in the German interest. Munich to Burgundy will probably fail as well, but if France has not supported the move then it has done enough. If France has supported the move then it is a good thing that the army is still in Munich, as it can now attack Burgundy to cut support, especially if Kiel moved to Holland, and so is now in a position to bounce in Belgium, without compromising the defence of Munich. And if the army in Burgundy moves away to Belgium then a German army replacing it threatens the French home centres.

Germany's long term threats are Russia and the Balkans but the game may be decided by the time Germany or the Balkans can mount an attack on the other. This means German strategy is based on crushing France and holding Russia in the North, until forces can be moved from the west.

7.1.5 Italy

Italy looks weakest, far from the neutrals in the Balkans, and in danger of being unable to take even Trieste by Fall 1901. The armies in Rome and Naples have a long way to travel to reach anywhere. Three home centres, three Austrian and three Balkan is only nine. Marseilles is can held by France indefinitely, so four are still required from Moscow, Germany and Russia.

The direct move from Venice to Trieste in Spring 1901 will probably bounce with Serbia. Moving Venice to Tyrolia, and Rome to Venice will allow a supported attack in the fall, but at the cost of allowing the army in Serbia into Trieste unopposed in the Spring. By the fall the army starting in Naples will reach the Adriatic, and Tyrolia can support Venice into Trieste or strike at Vienna, depending on the Balkan opening moves. However it may be no better than a guess.

An early strike on France is very difficult. It would require the army in Marseilles to be moving the wrong way, and France does not have any useful things for it to do that

will stop it from covering the gap south of Switzerland.

7.1.6 Russia

Russia will aim for the three home centres, Warsaw, three Scandinavian, three Balkan, and then three from Germany and Austria, probably Berlin, Vienna and Budapest. Kiel is a possibility instead of Serbia.

The Russian army in Sevastopol is a threat to the flank of the Balkans and has little to do other than march on Rumania. The army in St. Petersburg will move to Norway then to Sweden, possibly to bounce with German forces coming through Denmark, but capturing a centre regardless. Warsaw will be Russian again before too long, though it is actually stronger to play Moscow to Ukraine in Spring 1901, stopping an attack from the Balkans if it bounces, and forcing the Balkans to defend Rumania with support if it succeeds. Of course if Rumania is defended then the army (now in Ukraine) can still move to Warsaw for the fall, securing the second build. It is not nearly as dangerous for an enemy to occupy Ukraine that fall; with armies in Warsaw, Sevastopol and newly built in Moscow; any enemy can swiftly be beaten back. Also an army in Ukraine in fall 1901 cannot have taken a centre, with another defending Rumania that leaves the Balkan player with a single build.

After two builds Russia will look strong in Spring 1902. One army will join its comrades in Scandinavia, taking whichever of Norway and Sweden that remains neutral. While Germany could try to rush its builds north, Warsaw to Silesia or Prussia and Moscow to Warsaw will leave Germany fighting on three fronts, with five or possibly six armies facing two or three French to the west, and three or four Russian armies from the North and East. And if the Balkans or Italy secure northern Austria then Munich becomes tempting for them as well.

Alternatively Moscow could move to Ukraine in Spring 1902, along with Warsaw to Galicia. This clearly threatens Rumania in the fall and if captured from Sevastopol then a build is available in Sevastopol to help defend it. By Spring 1903 Budapest or Bulgaria may fall, and the rest of Austria and the Balkans are wide open.

Russia looks strongest, with the defensive advantage of a corner power, but lots of nearby neutral centres, but will have to fight both Germany and the Balkans from quite early on, and may not be able to fight off both to avoid defeat.

7.2 Five Player Results

In this test game (Game 2-6) the program suffered from being friendly with powers that it would have been better attacking, and pursuing some battles for longer than worthwhile. In addition it struggled to find a use for units far behind the front line, and had difficulty getting them into the front line when they finally got there. This test lasted until Fall 1917, but by 1906 the game was already decided.

7.2.1 Spring 1901

The Balkans moved Serbia to Budapest, and Bulgaria to Serbia, with Rumania holding. Not as strong as above, but still threatens Trieste in the fall. France moved Paris to Burgundy, with support from Marseilles, and Brest to Picardy. Germany moved Kiel to Denmark, Berlin to Kiel, and Munich bounced in Burgundy. Italy moved Rome north to

Tuscany, Venice to Trieste, and Naples held. Here Naples should have moved, anywhere would be an improvement, and even risking bouncing in Venice would have been no loss. The lack of any short term goal for units in Naples and Apulia would cause Italy to make numerous poor moves. Russia moved St Petersburg to Norway, Moscow to Warsaw, and Sevastopol bounced in Romania.

Overall Naples holding was the only useless move. The Balkans still can acquire two builds, but it could have guaranteed two regardless of Italy's move. And it can always build in Rumania, so bouncing twice would have been enough to defend it past the fall. Russia would have been stronger moving Moscow to Ukraine then Warsaw — this might be regarded as unduly provocative towards the Balkans, but war is almost inevitable, and anything that weakens the Balkans (even to the benefit of Italy) is probably good for Russia.

7.2.2 Fall 1901

The Balkans used Serbia to support Budapest into Trieste, and Rumania held again. France attacked Munich with Burgundy, Marseilles bounced in Piedmont, and Picardy moved into Belgium. Germany moved Kiel into Holland, Denmark bounced with Russia in Sweden, and Munich held. Italy used Tuscany to bounce with France in Piedmont, Naples held again, and Trieste tried to attack Budapest, but bounced and was dislodged into Vienna. Russia bounced with Germany in Sweden and with the Balkans in Romania, and held in Warsaw to secure the centre.

Again Naples holding was useless, and deciding Serbia was more important than Budapest was very poor (it is a threatened home centre, against a threatened normal centre, so is viewed as more valuable), as Budapest supporting Serbia into Trieste would have gained an extra build, ideal to defend Serbia. Every other move was sensible although Munich could have moved to Burgundy without loss.

Builds were made in Bulgaria, Paris, Kiel, Berlin, Venice, Moscow and St Petersburg.

7.2.3 Spring 1902

The Balkans, viewing Russia as the largest opponent, tried attacking Sevastopol from Rumania and moving Bulgaria into Rumania, which both bounced. Serbia supported Trieste, which was enough to defend against an Italian attack.

Italy finally moved Naples to Apulia, the move from Vienna to Trieste with support from Venice bounced and again bounced in Piedmont from Tuscany.

France moved Paris north to Picardy, and saw Belgium bounce with Holland, Burgundy with Munich and Marseilles with Tuscany in Piedmont. Germany attacked France from Holland and Munich, both bouncing, and moved Kiel to Ruhr. Against Russia it bounced in Sweden again from Denmark, and moved Berlin to Silesia, threatening Warsaw.

Russia again bounced with Germany in Sweden, attacking from Norway. St Petersburg moved to Finland to support next turn. In the south Sevastopol again bounced in its attack on Rumania, and Warsaw moved to Galicia and Moscow to Ukraine to add extra pressure.

Of the 22 moves, only Berlin to Silesia appears an outright mistake. The Balkans may have been better attacking Vienna, but could have only swapped it for Trieste. The army in Bulgaria is effectively trapped, but will now be needed to support Rumania.

France needs to take the Ruhr, as it is the only point that it can attack with support. It is slightly faster to move Paris to Burgundy and Burgundy to Ruhr than Paris to Picardy to Belgium. Germany needs to match the three French to the west, but also needs to match the two Russian units to the North, so the move from Berlin to Silesia will prove a mistake, as Scandinavia falls to the Russians and Germany can only bounce in Warsaw. Germany must secure Sweden or Belgium before striking east so it would have been better to move to the Baltic, then threaten Sweden, and possibly march into Livonia in a Spring turn, threatening three Russian centres in the Fall. This time Italy would have been better attacking Trieste from Venice, with support from Vienna. The danger is that Trieste retreats into the Adriatic then threatens the Italian home centres, but Italy has enough units to defend until the Winter, when Italy will have a build and the Balkans must disband a unit.

7.2.4 Fall 1902

The Balkans, now surrounded, held in Rumania and Trieste and supported Rumania with the armies in Serbia and Bulgaria.

France again bounced in Piedmont from Marseilles, Belgium to Holland, and Picardy to Belgium both bounced, and Burgundy to Munich bounced and was dislodged by a German attack It retreated to Paris .

Germany held with Holland, and used Munich to support an attack on Burgundy from Ruhr. Denmark bounced once again in Sweden and Silesia bounced when moving to Galicia, trying to threaten four centres instead of one.

Italy used Venice to support Tuscany into Piedmont, sealing the French off. Apulia continued into the Adriatic, and Vienna bounced in the still neutral Budapest.

Russia finally took Sweden from Norway, with support from Finland. Ukraine moved to defend Warsaw from the Germans, and Galicia bounced with Italy in Budapest. Sevastopol held, as the Balkans dropped off the list of enemies. Russia then used the extra centre to build in Moscow.

This was a good chance for the Balkans to take Budapest. The Russian attack was unlikely because Warsaw was threatened. Rumania and Trieste could have supported a move into Budapest from Serbia, or Trieste supports Rumania, with Bulgaria moving to cover Rumania. But Trieste was vulnerable to an Italian move, and Rumania was defended against an attack that was less likely. Venice supporting Tuscany into Piedmont looks good, but does not provide any advantage over France, and meant missing a chance to take Trieste and link with Vienna.

France again neglected the chance to strike at the Ruhr, allowing Germany into Burgundy but failing to move into the now empty Ruhr. Had it done so Germany would have to defend Kiel, but France could defend Paris, and either Picardy or Gascony, forcing Germany to guess, and with a good chance to take a German centre in the confusion. But equally Germany might have been better to move from Munich rather than Ruhr, keeping a better line against France, and not threatened by Italy. Then the army could swing into Belgium, with two support against one, producing the extra army to fight Russia, and maintaining a decisive advantage in the West. However an attack on Ruhr would have cut support, maintaining the status quo. The problem is that France values the two German centres higher than the Ruhr, and cannot bring a third unit into the attack.

Finally Russia has also supported with the wrong unit, as the army in Finland will

take several turns to move into a useful position unless it moves through Sweden, which it should have done now, then supported Norway into Skagerrak for the assault on Denmark. But Russia has still taken a centre, and only the French position looks any worse than last turn.

7.2.5 Spring 1903

The Balkans tried to move Rumania into Galicia, which bounced, and Bulgaria into Rumania, which then also bounced. To protect against a Russian attack they supported Bulgaria from Serbia. Trieste held, but was dislodged by Italy, and retreated to Albania.

France reoccupied Burgundy from Picardy, supported by Marseilles, Paris, and Belgium.

Germany held in Holland and Burgundy, with Burgundy being dislodged. Denmark attacked Russia in Sweden, Silesia attacked Russia in Galicia, both bouncing, and Munich bounced in Tyrolia. Burgundy retreated back to Ruhr. It would do more damage by moving to Gascony but at the cost of leaving a gap between Holland and Munich.

Italy moved Adriatic Plain into Trieste, with support from Vienna and Venice. Piedmont, no longer interested in France, tried moving to Tyrolia, but bounced with Munich.

Russia moved Finland to Bothnia, and Sweden bounced with Denmark. In the south, Galicia took the unoccupied Budapest, Sevastopol bounced again in Rumania, Warsaw bounced trying to move to Silesia. The new army in Moscow moved to Ukraine.

Germany is now concentrating on Russia, the biggest threat, and because Russia outnumbers Germany, Germany does not wish to be fighting on another front. This meant that it neglected trying something like Burgundy to Belgium, supported by Holland, choosing instead to hold with both. However if it had taken Belgium, France would retreat to Ruhr, and threaten Belgium, Kiel and Munich. If it was then an adjustments phase, this would be perfectly acceptable, as Germany can build in Kiel, and France has to disband. The problem is that here Germany wants to attack Russia, but does not realise of Russia's six armies, only three are in contact with German units, and two of those are in contact only because Germany has sent an army south east. Germany still has plenty of spare strength to use against France, and needs to crush France to free up units to invade Russia with to win.

Overall, France and Italy both attacked with every possible support, and both took ground. Russia took a centre. Germany failed to do anything in the West, losing an opportunity, and should, at least, tried Holland to Belgium, cutting support. Munich may have been better supporting Burgundy than trying to get involved with Austria and Italy. Austria suffered for its failure to take both Budapest and Trieste earlier, as Trieste is difficult to defend. The decision to stop attacking Sevastopol was good, but Budapest would have been a better alternative than Galicia. The biggest mistake was that Germany failed to do anything in the West, losing an opportunity, and should, at least, tried Holland to Belgium, cutting support. Munich may have been better supporting Burgundy than trying to get involved with Austria and Italy.

7.2.6 Fall 1903

The Balkans now go defensive again, with Serbia supporting Rumania, and everyone else supporting Serbia.

France also goes defensive, with Paris supporting Burgundy and Burgundy supporting Belgium. Belgium and Marseilles both held.

Germany held with Holland, and Munich supported Ruhr against Burgundy again. Denmark again bounced with Sweden, and Silesia advanced into Galicia.

Italy succeeded at the second attempt in moving Piedmont to Tyrolia, now supported by Venice. Trieste continued into Budapest, supported again by Vienna.

Russia attacked Trieste with Budapest, which failed and was dislodged and forced to disband. Sweden again bounced in Denmark and Sevastopol bounced in Rumania. Ukraine also bounced in Rumania. Warsaw advanced into Silesia and Bothnia into the Baltic Plain, putting lots of pressure on Germany in Denmark and Berlin.

So again, the Balkans fear attack from all sides and are paralysed. France should have moved Paris to Burgundy last turn, as can now no longer support Belgium, except from Burgundy which can easily be cut. Now Germany should attack Belgium from Ruhr, with support from Holland, possibly losing Ruhr, but guaranteed to take Belgium, gain a build and force France to disband a unit. If France had the army in Picardy instead of Paris it could cover both Burgundy and Belgium, as long as Italy did not attack Marseilles.

Italy successfully took Budapest, but Trieste is still controlled by the Balkans. Moving Venice to Trieste would have been more useful than supporting Piedmont to Tyrolia. Russia could do nothing to keep Budapest, but could have tried using Budapest to attack Rumania. But the biggest failure was both Ukraine and Sevastopol moving to Rumania.[1] However it has a substantial force building against Germany, and is in no trouble in the south.

In the winter, Russia rebuilt the army disbanded in Budapest in Moscow, and Italy built in Naples.

7.2.7 Spring 1904

The Balkans tried to move Albania into Trieste, which bounced. Serbia and Bulgaria both supported Rumania, which held.

France made no headway against Germany, with Belgium bouncing against Holland and Burgundy against Munich. Paris then bounced trying to move into Burgundy, supported to no effect by Marseilles.

Germany is now struggling. Holland has to defend Kiel from the Russian attack. Munich supports Ruhr to stop the French taking advantage, while Ruhr supports Munich against the risk of a combined Franco-Italian attack. Denmark supports Holland into Kiel, but the support is cut by the attack from Sweden. Galicia held.

Italy consolidates its gains last turn. Tyrolia supports Venice into Trieste, Naples moves to Apulia. Vienna turns to attack Galicia and Budapest attacks Rumania, both bouncing.

Russia tries to take Kiel from the Baltic, bouncing, but using Sweden and Silesia to cut support from Denmark and Munich respectively. Ukraine bounces in Galicia, and Moscow in Ukraine. Sevastopol supports Moscow into Ukraine, preventing the remote possibility of an attack. But Russia had no successful moves.

Germany is caught between three French and three Russian armies. Kiel is the most important province because it is adjacent to the four other German centres. Because of this, Germany defends it and Russia attacks. Russia could force Denmark or Berlin

[1] This bug has since been fixed, by checking for collisions at the start and end of the movement calculating routine.

instead, and should have, taking one then the other. The German army is Galicia is also adjacent to four centres, which makes Galicia more important than the centres themselves, so it held. This would be true if Germany had several units attacking, and it will be seen later that Galicia is the key to the south, but Germany should simply not be involved in this. The unit could be useful to win an alliance from Italy or the Balkans, but the program has no capacity for such strategy, and the unit is isolated and giving no help to Germany.

7.2.8 Fall 1904

Again the Balkans attacked Trieste from Albania with no support, which bounced. Serbia and Rumania mutually supported each other, and Bulgaria supported Rumania.

France again tried Burgundy to Munich, and supported Paris into Burgundy with Marseilles and Belgium, which both bounced.

Germany again tried moving Holland to Kiel with support from Denmark, Munich and Ruhr mutually supported each other, Galicia held again. But Russia attacked Denmark, taking it, and Germany disbanded the army.

Italy moved Apulia to the Adriatic, and attacked Rumania from Budapest, and Serbia from Trieste, which both bounced. Vienna to Budapest and Tyrolia to Trieste also bounced.

Russia attacked Denmark from the Baltic with support from Sweden. Silesia bounced in Munich and Ukraine bounced in Galicia, while Sevastopol held and Moscow moved to Warsaw.

Again France suffers because there is not enough width to bring its armies to bear. It must get an army into the Ruhr. From there it can swing round into Munich or Holland, move another army into Ruhr and Germany loses an army. This leaves Germany unable to fully defend itself against Russia and France. Munich scores higher than Ruhr, and Burgundy cannot possibly force Munich without help. The army trapped behind is doing nothing, but the program cannot do anything with it.

Italy has split its resources between Rumania and Serbia. Rumania has the higher score, so Budapest attacks it instead of supporting Trieste into Serbia. But as Sevastopol held, Italy cannot take Serbia anyway.

Russia has taken Denmark, but has again taken it with the wrong unit, and no longer threatens the supported attack on Berlin. Sweden will have to move to the Baltic next turn.

The Balkans disbanded Bulgaria, Italy built in Venice and Russia in Moscow.

7.2.9 Spring 1905

The Balkans tried Rumania to Galicia, then Serbia to Rumania and Albania to Serbia. They cannot risk losing a centre. However it suggests that the decision to disband in Bulgaria rather than Albania was a mistake, as it would be easier to cover Rumania from Bulgaria and hold or support with Serbia. As Russia is now the largest power, the Balkans are trying to attack Russia so ignore Trieste and Serbia.

France again tries Paris to Burgundy and Burgundy to Munich. But Germany has left Holland to defend against Russia, and so Belgium takes Holland unopposed.

Germany used Kiel and Ruhr to support Munich, and Munich supported Ruhr. Galicia again held, but was dislodged by Russia, and retreated to Bohemia.

Italy again attacked Rumania from Budapest, but tried to follow it with Vienna to Budapest, supported by Trieste, and Tyrolia to Vienna. All three bounced. Adriatic and Venice both supported Trieste. Again Italy must move against Serbia instead of Rumania, and should move an army into Albania to provide enough support.

Russia attacked the German army in Galicia from Warsaw, supported by Ukraine. Moscow supported Ukraine, and Sevastopol held. Silesia bounced in Munich, Denmark in Kiel, and Sweden moved into the Baltic Plain.

7.2.10 Fall 1905

The Balkans supported Rumania with Serbia, and Serbia with Rumania and Albania.

France again tried Burgundy to Munich, and Marseilles again supported Paris to Burgundy, with no result. The army in Holland bounced in Kiel.

Germany supported Kiel with Munich and Ruhr, and Munich with Kiel. Bohemia tried to move back to Galicia, but failed, and was disbanded in the winter. Having lost Holland and Denmark, with Berlin undefended, and three French armies and three Russian armies on surrounding the German home centres, Germany has little prospect for capturing any centre, and will spend the rest of the game mutually supporting, trying to survive as long as possible.

Italy is now concentrating purely on Russia, as the highest scoring power, in contact with an Italian unit, and outnumbering Italy. As it cannot attack the Balkans, all the Italian units support a neighbour. Only Budapest was attacked by another power. Italy has to outnumber a Balkan unit to stab it, but it will never manage this as it can only bring three units to bear against Serbia, and there will always be three enemy units in and adjacent to Serbia

Russia attacked Rumania from Sevastopol with support from Ukraine, but could only bounce. Galicia bounced in Budapest, Silesia in Bohemia, Moscow moved to Warsaw, and Denmark and Baltic both held.

7.2.11 Test Conclusions

By the end of 1905, Germany is almost out, but it will not be eliminated until 1913. Russia takes all the German centres. The strategy algorithm for the losing powers forces them to attack Russia, and the game grinds to a draw. France has a stalemate line through Holland, Ruhr, Burgundy, Switzerland and Marseilles, although it does not properly defend it. But Russia is attacking the strongest and weakest powers — Italy and Austria — so does not and will never attempt to attack the French units. But Russia has enough units to stop France from ever taking Kiel or Munich.

Similarly, Austria will never take Sevastopol before Ukraine, and cannot even take Galicia. But Russia is a bigger threat than Italy, and so Austria will only ever attack Russia.

Italy holds the line Tyrolia — Bohemia — Vienna — Budapest, and extends it to include Galicia. But Russia holds the other side, and neither can make, and hold, a breakthrough

In comparison, by 1906 in game 4, Italy had attacked France through an undefended Marseilles, and captured all three French home centre with a single unit. Austria could gain no improvement on four units, and Russia could not reach Italy for Germany and

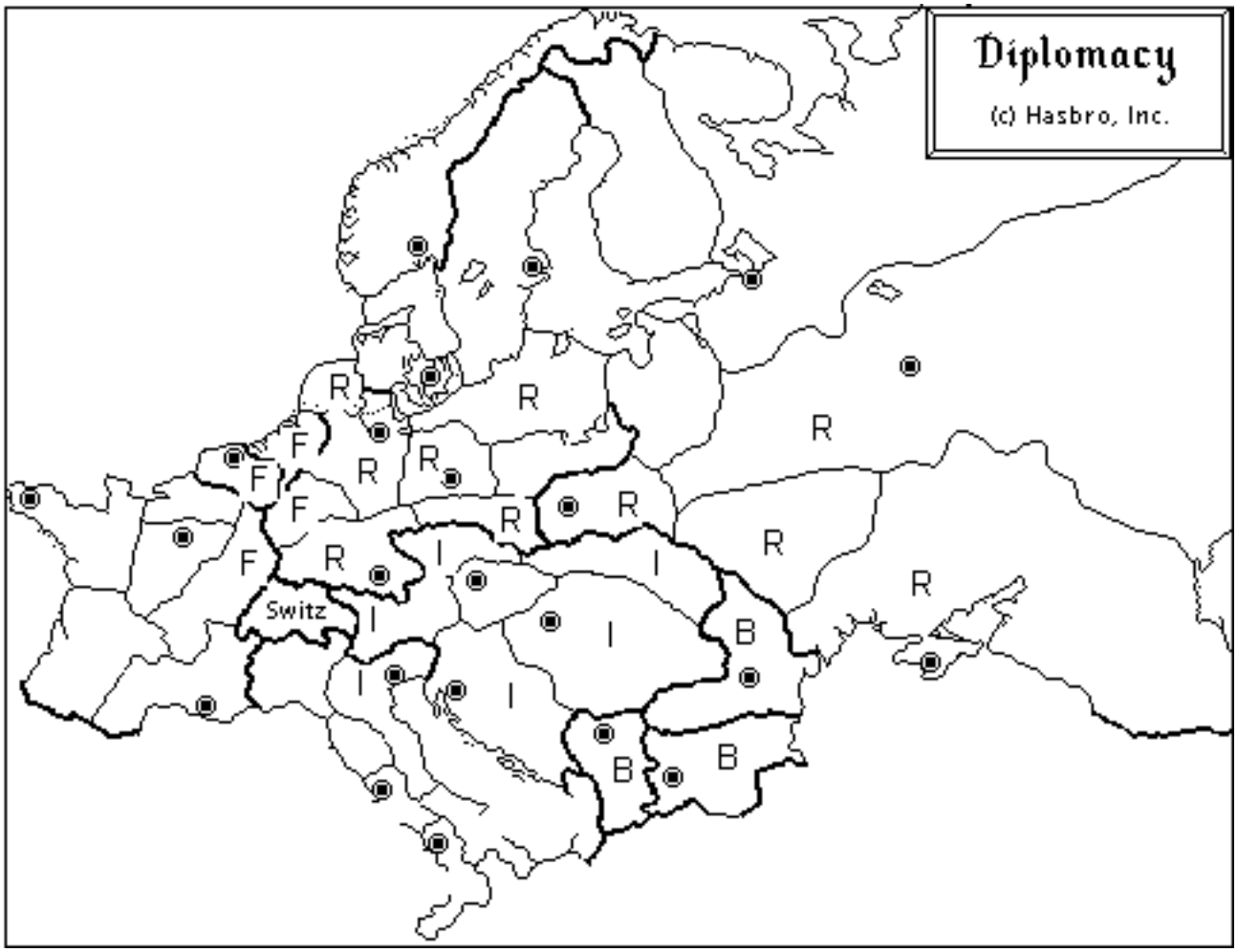

Figure 7.2: The final position — 4 way draw

Austria, but Germany could not attack Italy in France for fear of the Russian units on her borders.

And in game 5, Germany was reduced to a single centre, but France, Italy and Russia were too involved with each other to eliminate the final German unit.

The first conclusion is that the tactics algorithms work the vast majority of the time. However the location scoring occasionally sends them to the wrong place. Germany lost when it moved from Berlin into Silesia then Galicia, with no benefit to itself or its allies. A superior move was found in another test when it moved north, supported a bounce in Sweden, then moved into Livonia, threatening three Russian centres. But the disruption was not enough to cripple Russia (despite a German attack on Sevastopol from Moscow in Spring 1903).

The other instance was when Russia took Sweden and then Denmark with the wrong unit. The problem is that the supply centre looks threatened, so the army in it supports rather than attack, although it cannot possibly lose the centre in that turn.

The strategy is overly simple, but kept most of the powers in the game. Perfect Diplomacy should see the leader attacked and dragged down until a new leader appears, and another coalition forms against them. It is not quite what is happening, but the result is the same, a draw split between most of the players.

Chapter 8

Further Work

Firstly there is some work still to be done before fleets are properly handled. The results are parsed properly, but currently a fleet cannot convoy an army, fleet movement to and from bicoastal provinces does not specify a coast, and fleets cannot be built. The first is an addition to the movement code, allowing fleets to convoy an adjacent army instead of moving themselves. The second is largely presentation, although there will be a problem when a fleet has a choice of coasts to move to. The third really requires some knowledge of what the build would be used for. Instead the next best solution is to build according to a ratio of armies to fleets, and this could be modified depending on the strength of the nation currently being attacked by the one with the build. The scoring will have to be modified as well for fleets to be properly used, currently sea areas score very low compared to land.

Secondly, the program has several rough edges. The input checking could be improved as wrong input will cause the program to crash, e.g. units that were not moving being marked as bouncing or two units that have not been dislodged occupying the same location. The display of the current season can also be misleading as it is the next season that the program expects to produce orders for, not the next season that it expects results for. It will currently display that it is a movement season after it has produced retreat moves, but before it has read the retreat results. This is because the retreating units are destroyed, so the program no longer has a record of the retreating units.

It was intended that the program would be run when judge results were received. In testing the program was used to run through an entire game in a single session. The program does have a limited load/save facility which covers ownership of locations, and the units are placed from the judge input. However each power's ratings of its rivals are not stored. These should be added to the saved game file to allow a game to be run over a longer period.

Diplomacy has many variants, and it was hoped to make the program as generic as possible, to allow different variants to be run. Hence the use of unmodified judge map files to allow the program to run with different maps. However the number and names of the powers are currently fixed in the code. These could be removed to a variant file, allowing the program to run variants using different sets of powers as well as maps. Starting locations are already determined by the first judge result file. If the variant file is then used to create the powers in each game, it could be combined with the save game file to include names and starting strategy ratings for each opponent for each power.

The problem with the program tactically is that it is incapable of forward planning. This is most obvious when a human expects an attack to bounce, and would look to move

an additional unit to support an attack in the next turn. The *location* scoring algorithm takes potential moves into account, and the *nation* movement algorithm looks at chains of moves, and tries to find an alternative if if the first move bouncing will cause the others to bounce as well. But neither is as useful as even a short search, that would reveal that the delaying the attack for support is more likely to be successful.

The branching factor is very high, as described previously, but by only including units within say 3 spaces of the searching power's units, and limiting it to the next two turns, it might be kept to a manageable level. The other problem is assessing the probability of enemy moves, but the existing algorithm could be used to provide an estimated probability distribution of enemy moves. It would also allow the simplification of the location scores, with a move's utility depending on supply centres captured, and units in contact with the enemy. Finally it would reject many pointless moves, such as units bouncing with friendly units that were holding.

By running several short searches, it would perhaps be possible to take probabilities for the best moves for each nation, and use them in the next search for the other nations. Ultimately the moves could be randomly generated from the set of probable moves, getting past the difficulty that there is not always a best move, merely set a moves that may win against some enemy moves, and lose against others.

Strategically there is no ability to support a unit of another power. This prevents true cooperation between powers, and reduces them to mounting separate offensives against a common opponent. However it does require some idea of what the ally intends to do. In no-press games, this is achieved by using support and convoy orders, intended to fail, which pass information to the other players. Fleets convoy a unit of another nation to an enemy capital are perfectly legal orders syntactically, but cannot possibly succeed, especially if either the unit being convoyed or the destination is land-locked. An ability to support allied units, or at least produce message orders and understand message orders as such, would be useful, and is the next step to improve the standard of play from the program.

After that the next logical strategic step would be to move to limit-press diplomacy, and allow the program to pass messages to other players to produce alliances, split neutral centres in a mutually satisfactory fashion, and request support for moves.

The program includes many arbitrarily valued constants. The location scoring values were hand optimised by looking at the poor moves made in the two player tests and increasing the appropriate constants. However it was difficult to remove all weaknesses, and it led to value inflation. There are many more rigourous techniques available which were rejected because of time considerations, but would result in an increased performance.

An automated Judge-style adjudication program would be useful and the Judge algorithm is freely available. It could be combined (possibly with the addition of command line parameters to the A.I. program) to enable the A.I. program to load results, play a turn, output results, have them adjudicated, and repeat until the game ends. This would allow a far faster testing pace than was achieved, at the cost of a significant amount of effort to set it up. The program could then be tested against other versions of itself, using genetic algorithms to optimise the constants.

Chapter 9

Conclusion

The program plays the two player version strongly, although Austria makes enough mistakes to lose against a better opponent. The five player version is a lot weaker, as the program cannot see the overall picture and act accordingly.

The five player version finds good moves, but often takes several turns to do so. Often a stalemate is reached until a unit reaches a key area (Ruhr, Helgoland or Albania in the example given), when a human would act to reach the area as fast as possible.

A difficulty was that the five player map was lacking the breadth for a mobile campaign, and several turns would pass before someone moved. This would be fixed by the fleets in the seven player game, and it appears better to ignore this problem and concentrate on the full game.

Another problem with the five player version is that a successful power quickly comes into contact, and is attacked by, three opponents, resulting in the stalemated positions. But again this is not an issue in the full game.

The tests with judge input were successful, but the output is not complete (the bi-coastal problem), and the two and five player games were all hand adjudicated.

The test games took too long to play, and this hindered the optimising of coefficients. There is scope for improvement using an automated adjudicator, and comparison against human games and other optimisation techniques.

Overall the tactical problems have been solved. The strategic problems are more difficult, but implementation of fake orders to pass messages will make cooperation easier.

Bibliography

[1] Dave Hemke. Reviews of microprose pc diplomacy. *www.diplomacy-archive.com*, 1999.

[2] Danny Loeb. Challenges in multi-player gaming by computers. *http://devel.diplom.org/Zine/S1995M/Loeb/Project.html*.

[3] Sarit Kraus and Daniel Lehmann. Designing and building a negotiating automated agent. *Computational Intelligence*, 11:132–171, 1995.

[4] Allan B Calhamer. *Rules of Diplomacy 4th ed.* Hasbro, 1999.

[5] Richard Sharp. *The Game of Diplomacy.* 1978.

[6] T A Marsland. The anatomy of chess programs. *http://www.cs.ualberta.ca/ tony/ICCA/anatomy.html*.

[7] Stuart Russell and Peter Norvig. *Artificial Intelligence, A Modern Approach.* Prentice Hall, 1995.

[8] Darse Billings, Aaron Davidson, Jonathan Schaeffer, and Duane Szafron. The challenge of poker. *Artificial Intelligence Journal*, 134:201–240, 2002.