Advanced Computer Recognition of Aesthetics in the Game of Chess

AZLAN IQBAL and MASHKURI YAACOB
College of Information Technology
Tenaga Nasional University
Km 7, Jalan Kajang – Puchong, 43009 Kajang, Selangor
MALAYSIA
azlan@uniten.edu.my, mashkuri@uniten.edu.my

Abstract: - This research intended to see if aesthetics within the game of chess could be formalized for computer recognition since it is often appreciated and sought after by human players and problem composers. For this purpose, Western or International chess was chosen because there is a strong body of literature on the subject, including its aesthetic aspect. Eight principles of aesthetics and ten themes were identified. Flexible and dynamic formalizations were derived for each one and cumulatively represent the aesthetic score for a move combination. A computer program that incorporated the formalizations was developed for testing purposes. Experiments were then performed comparing sets of thousands of chess compositions (where aesthetics is generally more prominent) and regular games (where it is not). The results suggest that computers can recognize beauty in the game. Possible applications of this research include more versatile chess database search engines, more accurate automatic chess problem composers, enhanced automatic chess game commentators and computational aid to judges of composition and brilliancy tournaments. In addition, the methodology applied here can be used to gauge aesthetics in similarly complex games such as go and generally to develop better game heuristics.

Key-Words: - aesthetics, chess, game, evaluation, intelligence, computation

1 Introduction

In the game of chess, aesthetics is important and appreciated not only by grandmasters but average players as well. Garry Kasparov, arguably the world's strongest player is reported to have said, "*I want to win, I want to beat everyone, but I want to do it in style!*"[1]. Computers currently play chess at grandmaster level and have even defeated the world champion but they cannot tell an attractive or beautiful combination from a bland one because the objective has always been simply to win [2-4].

This is also why computers have been unable to create or compose chess problems like humans do. There is a sufficient body of literature on chess that adequately covers its aesthetic aspect (refer section 2) and the research presented here was intended to see if this information could be formalized for computational purposes. The result is a model of aesthetics that consists of unique formalizations of the principles of beauty in chess, which includes several themes. It is potentially capable of giving computers the ability to recognize aesthetics in the game like humans do.

Section 2 reviews some of the important contributions to the area. Section 3 details the proposed formalizations and Section 4 presents some experimental results intended to validate them.

A discussion on the results and related issues appears in section 5. The paper concludes with a summary and suggestions for further work. With over 700 million chess players and composers worldwide, the authors believe this research presents significant findings with respect to AI within the domain of chess itself even though extensions to other games or areas are not fully explored in this paper [5]. However, a brief discussion on such extensions is presented in section 5.1. The information that follows is therefore specific to chess - as it is required to be for efficacy - given the inextricable nature of aesthetics to its domain.

2 Review

One of the earliest formal references to the aesthetics of chess was by former world champion Emanuel Lasker in his book, "Lasker's Manual of Chess" where he devoted an entire chapter to it. There he writes of the concept of "achievement" (e.g. winning material, space, the game itself) being important to aesthetics and that understanding of the game, not mastery, is all that is required for its appreciation [6]. Margulies, a psychologist, derived experimentally eight principles of beauty in the game from the judgement of experienced players, as follows [7].

1. successfully violate heuristics
2. use the weakest piece possible
3. use all of the piece's power
4. give more aesthetic weight to critical squares
5. use one giant piece in place of several minor ones
6. employ themes
7. avoid bland stereotypy
8. neither strangeness nor difficulty produces beauty

Similar criteria have been mentioned in other sources [8-10]. Levitt and Friedgood add the notable concepts of geometry (e.g. graphic effects such as alphabets formed on the board) and flow (i.e. forced play rather than many confusing alternate variations) as additional elements of beauty in the game [10]. Aesthetics is not limited to compositions and also occurs in real games, though less often [12-14]. Brilliancy prizes are even awarded at certain tournaments to games that are aesthetically noteworthy either in full or part [15].

Even though not all composition conventions (i.e. general practices) apply to real games, aesthetics is shared between the two domains as long as the rules are the same. Given say, direct-mate compositions (mate in *n* moves against any defense) they only differ with real games in terms of perceived beauty. Experienced players can often easily tell if a position looks like a composition because it is too "unusual" or "convenient" to have occurred in a real game [16].

Computationally, aesthetics has been left largely to humans since computers are capable of deriving forced checkmates by constructing a complete database (e.g. from a set of desired pieces) and working backwards one ply (half-move) at a time but not capable of any "creative" activity [17][18]. It is left to humans to judge if the constructed problems are beautiful despite being conventionally "correct" from a composition standpoint. Composition conventions (e.g. *include variations*, *no duals*, *no symmetry* etc.) are often used to benchmark chess problems computationally with little emphasis on aesthetic factors [19][20]. The two sometimes overlap in part but are usually distinct concepts. Real games for example, also exhibit aesthetic properties but do not adhere to most composition conventions (usually in excess of 20 "rules") [21].

Previously, only chess themes (e.g. Grimshaw, Pickaniny, direct battery), as a principle of beauty, had been weighted for the purpose of automatic chess problem composition and this was done by consulting one or two master composers [19][20]. The values ascribed to themes (especially exotic ones used in chess compositions and seldom in real games) were arbitrary and based on experience. This meant that some themes were preferred over others and that some or all themes might have to be weighted again if new ones were added since their values were relative to one another. Additionally, all implementations of a particular theme were therefore valued equally even though some configurations would no doubt be more beautiful than others [22].

Walls showed that beauty principles performed better than regular chess heuristics in solving certain types of chess problems [23]. He combined and incorporated a selection of Margulies' principles but used them to guide the game playing engine instead of evaluating the principles themselves so they were merely identified computationally as either being present or absent in a particular line of play. Hence, in terms of say, distance (or using all of a piece's power), a queen moving a certain number of squares across the board was considered just as "beautiful" as a rook or bishop.

For this research, weighting individual principles through supervised or unsupervised learning was not suitable because reliable test data (i.e. aesthetically rated positions) is scarce and more importantly, do not account for varying implementations of a particular principle [24]. It was also unnecessary since chess is a limited and precise domain with its own established measures and units that are not subject to personal taste in the way that say images are. In the latter case, linear regression or classification can be used to individually weight aesthetic features since there are no agreed standards for rating them [25].

The approach taken by this research is more akin to how the aesthetics of music is sometimes calculated, where discrete representations (e.g. frequency of notes, intervals etc.) of particular attributes (e.g. pitch, volume etc.) are used to recognize beautiful music [26][27]. However, chess is a more limited and less culturally-dependent domain than music so formalizations based on established metrics are probably more reliable. The next section describes in detail the metrics, chosen principles and scope of the research.

3 Methodology

In 1950, Claude Shannon explained how a computer could be programmed to play chess using estimated values of the chess pieces (K=200, Q=9, R=5, B/N=3, P=1) so that a score for every position in the game tree could be obtained based on the amount of material captured [4]. The king is essentially of ‘infinite’ value since its capture means losing the game but for practical programming purposes, it is often valued significantly higher than all other pieces combined. Using this method, computers could then decide which moves were the most favourable from a material standpoint and play a reasonable game of chess. Modern chess programs essentially still use Shannon’s methods with material being a primary factor for evaluation, even though piece values are sometimes changed during the course of a game given positional considerations [3].

For this research, the standard Shannon piece values were used except that the king was set to 10 since in aesthetics, winning is basically a prerequisite and there is no intention to drive game play. Additionally, “mating squares” or squares onto which occupation by the attacking piece would result in checkmate are also legitimate threats and valued equivalent to the king. Aesthetic evaluation of a chess combination is performed in retrospect on the completed move sequence to determine how beautiful it is. The squares of the chessboard itself are used as a metric to evaluate properties like distance and piece power because more powerful pieces tend to control more squares [28]. Distance is measured as the number of squares between two pieces on any line (i.e. ranks, files or diagonals).

If there are three squares between two pieces, the distance is calculated as four; starting from the location of the first piece and moving one square at a time, ending at the location of the second. Piece power (i.e. mobility) refers to the maximum number of squares a piece could possibly control on an empty board and was found to be: king (8), queen (27), rook (14), bishop (13), knight (8) and pawn (4). The pawn’s power is based on the fact that it can capture one square to the left or right and move forward one square or two for a total of four. Piece power is used to attribute slightly different values to identical maneuvers performed by different pieces. It is based on their relative importance as generally perceived in the game.

3.1 Selected Principles and Scope

Based on the literature surveyed (refer section 2), eight aesthetic principles in chess were identified and selected namely, violate heuristics, use the weakest piece possible (to checkmate), use all of the piece’s power, win with less material, sacrifice material, economy, sparsity and use of themes. Margulies’ 4th principle was not explicitly included because it simply means to emphasize the role of the “active” (i.e. moving or checkmating) piece in a move sequence. His 5th principle used imaginary pieces not within the scope of Western chess while the 7th and 8th principles rely on previous knowledge and experience so they could not be included. Winning with less material and sacrificing material is considered paradoxical and therefore aesthetic [8][10][11].

Geometry was not included because it is extremely rare, even in compositions while flow tends to be biased against compositions that typically feature many side variations and are even lauded for it. The goal of this research was to evaluate aesthetics in both domains (real games and compositions) but only where it is equally applicable. Given the variety of chess and feasibility issues, aesthetic evaluation was limited to mate-in-3 move sequences. This permitted access to a wide selection of chess compositions and combinations in tournament games. Each principle was also designed to score a maximum value of approximately 1 so there would be no arbitrary preference given to any. There was nothing in the literature surveyed to suggest that some principles are inherently better than others. For explanatory purposes, white is always assumed to be the winning side.

Checkmates - though preferably forced (like in direct-mate problems) - are also considered aesthetic even if they are not forced. A beautiful checkmate combination in a real game for example, is often due to the oversight of the opponent. It might be perceived by humans as “less beautiful” but only upon closer analysis and this would have little to do with the beauty of the actual maneuver played [15]. A composition however, would be considered invalidated under these circumstances if it was of the direct-mate variety but this has more to do with convention (i.e. it must be “correct”) than aesthetics. Self-mate problems for example, require that both sides cooperate to checkmate black, primarily because certain (aesthetic) effects are not possible with direct-mates [29]. The selected principles and the rationale behind their proposed formalizations are explained in the following subsections.

3.1.1 Violate Heuristics

Heuristics in chess are essentially rules that govern good play. A move that violates one or more heuristics is considered paradoxical if it results in an achievement of some kind (e.g. checkmate). Given the scope of mate-in-3, four heuristics were selected for evaluation: keep your king safe, capture enemy material, do not leave your own pieces *en prise* (i.e. in a position to be captured) and increase mobility of your pieces. Other less tangible heuristics such as control the center and avoid doubled pawns were not included [30][31].

A violation of keep your king safe was defined as moving the king to a square which makes it prone to check on the next move. If the king's destination is in the center four squares of the chessboard, it counts as a full violation and scores 1 full point. If in the immediate outer 12 squares, 0.75 points. The next surrounding 20 squares, 0.5 points and the edge of the board, 0.25. This is because there is greater risk of exposure as the king approaches the center of the board.

Not capturing enemy pieces that are exposed and could be captured advantageously counts as a violation. Given the complexity of some exchanges and related positional dynamics in chess, only undefended pieces or defended ones worth more than the capturing piece qualify. A non-capturing move or one that prefers a different piece than the most valuable available violates this heuristic. Pawns are not included as pieces worth capturing because they are not valuable enough to get sidetracked and fall short of what is required for a decisive advantage in chess (i.e. 1.5 pawns) [32-34]. The score for this principle is calculated as the sum of the value of uncaptured enemy pieces divided by the value of the queen. Therefore a full point is scored in cases where a queen or pieces of equivalent worth are not captured in favour of some other move.

Like the previous violation, leaving your own pieces in a position to be captured (or *en prise*) applies only to pieces and not pawns. There is no violation if the move played captures an enemy piece worth more than the one left en prise or if the friendly piece is favourably defended (no potential loss of material). The score is calculated as the sum of the value of *en prise* pieces divided by the value of the queen.

The last violation is decreasing your own mobility. Usually, players try to control more squares with their pieces but sometimes the opposite is done and this can be quite obvious and puzzling. For example, a queen or bishop may be moved to the very corner of the board behind some friendly pieces where its mobility is greatly reduced or moved to block several other pieces, reducing general mobility. The score is calculated as: $(w_1-w_2)/w_1$; $w_2 \leq w_1$. Here, w_1 is the number of possible moves for white in the initial position and w_2 is the number of moves after his first move (assuming for a moment, white still had the move). Violation occurs only if the result is a positive value.

Heuristic violations are determined only after white's first or key move because in compositions the first move is usually enough to solve the problem and by convention, the most surprising to solvers. Other moves in the sequence may exhibit similar characteristics but the paradoxical effect is not the same. The overall score for the principle of violating heuristics (P_1) is formalized as follows:

$$P_1 = \frac{\sum_1^n v(h_n)}{n}$$

$v(h_n)$ = value of a particular heuristic violation

(1)

3.1.2 Use the Weakest Piece Possible

This principle simply means using the weakest piece possible to achieve a particular objective. Given the scope, it was extended to mean using the weakest piece possible *to checkmate* and therefore applies to the last move in the combination. The score increases as the piece power of the checkmating piece decreases. The formalization is given as:

$$P_2 = \frac{4}{r(p)}$$

$r(p)$ = piece power

(2)

The numerator is set at 4 so that if the weakest piece on the board (i.e. the pawn) is used to checkmate, the score reaches its maximum of 1. In the case of a double checkmate (two pieces attacking the king simultaneously with mate), only the piece that moved (i.e. the critical piece) counts.

3.1.3 Use all of the Piece's Power

Using all of the piece's power is related to its maneuverability and can be interpreted as the number of squares a piece traverses in a single move. Traveling a greater distance is considered more beautiful than a shorter one. If a less powerful piece (e.g. bishop) travels a certain distance, more of its total power is used than if a more powerful piece (e.g. queen) travels the same distance.

Therefore the bishop move is considered more beautiful than the queen move. This principle applies to all moves of the winning side in the move sequence. The opponent's moves are not included because they usually work against the desired achievement (and hence aesthetics) of the winning side. The score is calculated as follows.

$$P_3 = \sum_1^n \frac{d(p_n)}{r(p_n)}$$

$d(p_n)$ = distance (in squares) traveled by a piece, $r(p_n)$ = that piece's power, n = number of evaluation stages (i.e. each move by white + checkmate) (3)

The knight, given its unique movement defaults to a fixed number of squares (i.e. 3). The average maximum piece power of the chessmen is 0.29. In a mating combination, the distance between the checkmating piece and the enemy king (after the final move) is also evaluated by this principle. It is possible in certain positions for the total score to exceed 1 (e.g. two maximal pawn moves + one knight move + mate using knight = 1.75) or fall significantly below it (e.g. two single square queen moves + one single square rook move + mate using rook right next to the king = 0.22).

Like the previous principle, it applies to all combinations regardless of how beautiful or bland they might be so a deviation from the intended 'principle value limit' of 1 does not give this principle preference over others that could be absent in some combinations.

3.1.4 Winning with Less Material

This principle is considered aesthetic because it is paradoxical. Usually, the side with more material is more likely to win. It applies only if black's total material worth exceeds white's. The value is calculated as:

$$P_4 = \frac{(b_1 - w_1)}{m}, b_1 > w_1$$

w_1/b_1 = initial material of white/black, m = 38 (4)

The denominator is set at 38 because this is the maximum amount of expendable material for an army (at least one pawn must be left) where checkmate is still possible, however unlikely. The score is calculated at the initial position, prior to any moves.

3.1.5 Sacrifice Material

Sacrificing material is also paradoxical. It is not exactly the same as violating the heuristic of leaving your own pieces *en prise* because it applies more to exchanging your pieces unfavourably for positional superiority that is enough to secure a decisive advantage or force a win. The "romantic" players in the late 18th and early 19th century often used bold sacrifices that were not always sound to impress spectators [35]. Former world chess champion Mikhail Tal, who considered chess first and foremost an art, was also known for intuitive sacrifices that gave rise to complications on the board and confused his opponents [36].

In this day and age however, sacrifices are not as popular in real games or compositions because computer analysis can easily reveal flaws in them. Even so, sacrifices are still employed - even required in some positions - but are more calculated and scrutinized than before.

$$P_5 = \frac{w_1 - w_2 - b_1 + b_2}{m}, m \in \{9, 14, 19...\}$$

w_1/w_2 = initial/final material of white, b_1/b_2 = initial/final material of black, m = material constant (5)

The "dramatic effect" of a sacrifice usually correlates with the amount of material lost so the function above is used to calculate the value for this principle. The material constant consists of a set of values depending on how many moves there are in the combination. For example, a mate-in-2 sequence would have a material constant of just 9 because this (a queen) is the most amount of material that could be lost to the opponent in that short 'time'. A mate-in-3 would have a constant of 14 since after the opponent's second move, at most another rook (given the original piece set) could be lost and so forth. No sacrifices are possible for mate-in-1 positions and only positive values apply.

This function takes into account sacrifices of any number of pieces of any type, including adjustments for pawn promotions by both sides because the nett difference in material at the end of the move sequence will reflect how much material was really lost. It would be misleading for example, to sacrifice a knight after the first move only to promote a pawn to a queen on the second. Negative values indicate that white actually gained material but this is not held against him because many mating combinations necessarily result in significant material loss by the opponent. They are however, less beautiful.

3.1.6 Economy

Economy refers to using the minimal amount of resources to achieve a particular objective. For the scope, the objective is to checkmate the opponent. This principle is therefore evaluated in the final position where economy is most often exemplified [37]. It is difficult to ascertain economy in the moves preceding the final position because they may contain sacrifices or "quiet" maneuvers that are necessary but do not make much use of a piece's power. Also, the objectives of those individual moves are not as clear as it is in the position that results after the final move. Economy can be formalized as:

$$P_6 = \frac{\sum_1^n \frac{a_n}{f_n} - \left(\frac{o + \sum_1^n s_n}{f_k} \right)}{p}$$

a_n = active control field of a particular useful piece, f_n = maximum control field of that useful piece , o = overlapping control field square, f_k = maximum control field in king's domain (i.e. 9), s_n = maximum control field of a superfluous piece, p = number of friendly pieces on the board (including king)
(6)

The features here are derived essentially from the conventions employed in Bohemian problems which are known for their emphasis on economy [38]. A detailed explanation of this function and all the parameters can be found in [39].

3.1.7 Sparsity

Positions that are cluttered are considered less beautiful than those more spaced out [9]. An important feature when evaluating sparsity is therefore the number of pieces on the board in relation to available space. Even so, a position that requires more pieces should not necessarily suffer in terms of being sparse than say, an endgame position where pieces are inherently few. There are many ways that sparsity or its inverse, density can be evaluated (e.g. like pixels in a matrix, using quadrant density ratios etc.) but they do not translate as well to the chessboard [40]. For instance, a relatively 'dense' quadrant of the chessboard may be considered sparse if there are only 3 or 4 pieces because it is not practical or useful for them to be evenly spaced in different corners of the board.

There are also complications when we consider the centre 4x4 squares of the board as constituting a 'fifth' quadrant because sometimes pieces are concentrated there. In fact, activity or checkmates at the center of the board are considered more beautiful than at the edge or corner [41]. A more effective method to determine sparsity that works well with chess (and other similar board games) was developed and shown below.

$$P_7 = \frac{1}{\left(\frac{\sum_1^n \sum s(p_n)}{n} \right) + 1}$$

$s(p_n)$ = pieces surrounding a particular piece
(7)

Surrounding pieces are those in the field of a particular piece (i.e. immediate squares around it). Fewer pieces around a particular piece make the area around it appear sparser. The field is used because if there are say, only four pieces on the board they are considered sufficiently distant from each other (or not perceived as densely packed) even with only one square between them.

The average number of surrounding pieces is used to provide a better general idea of how uncluttered a position is. One is added to the denominator to prevent a division by zero error where there are no surrounding pieces around any of the pieces. Both black and white pieces are taken into account. Given that mating combinations often require the attacking pieces to be in close proximity to the enemy king, this principle is evaluated only in the initial position before any moves are made.

	
a) K. Fabel, Deutsche Schachzeitung, 1st Prize, 1965; sparsity: 0.264	b) Mannion vs Rojas, Yerevan ol (Men), ½-½, 1996; sparsity: 0.478

Fig. 1 Sparsity evaluation of chess positions

Figures 1a and 1b show the sparsity values of positions taken from a composition and real game respectively. The higher the score, the sparser it is. Simply adding or removing pieces will not necessarily bring the score down or up. It depends on where new pieces are placed and which existing ones are removed. Evaluations of many different positions suggested that this function captured the general perception of sparsity in chess better than alternative methods.

Themes in chess are essentially good tactics. Common themes include the fork, pin and skewer whereas more exotic ones, used primarily in chess problems, include the Grimshaw, Pickaniny and Plachutta. The effective use of themes is fundamental to aesthetics in chess. Only themes that are common to both compositions and real games were selected [42][43]. The themes and formalizations selected were the fork (T_1), pin (T_2), skewer (T_3), x-ray (T_4), discovered/double attack (T_5), zugzwang (T_6), smothered mate (T_7), crosscheck (T_8), promotion (T_9) and switchback (T_{10}) [11][43]. In the following equations, *d()* denotes the Chebyshev distance between two pieces and *r()* the piece power.

3.1.8 The Fork

The following evaluation function is used to evaluate the fork.

$$T_1 = \frac{1}{f_c} \cdot \left[\left(\begin{array}{c} \sum_1^n v\left(fp_n\right) + n \\ + \left(\sum_1^n \frac{d\left(f_k, fp_n\right)}{r\left(f_k\right)} \right) \end{array} \right) - k \right]$$

f_c = fork constant (i.e. 37), fp = forked piece, f_k = forking piece, k = number of checking moves by fp

(8)

The benchmark of the fork, *fc* was determined by first selecting the average number of possibly forked pieces (i.e. between 2 and 8) which is 5. The value of the most valuable pieces on the chessboard that could be forked in that way (assuming only the original set of pieces) namely the king, queen, two rooks and a bishop was then summed and added to the corresponding number of prongs required. The latter is equivalent to the number of forked pieces, *n*.

Fig. 2 Fork position involving mating square

The absolute maximum of 8 forked pieces was not used because this is extremely unlikely. Benchmarks should be reasonable. Possible checking moves, *k* by the forked pieces and intervening ones (assuming there are any) are considered liabilities and subtracted from the total. One of the peculiarities in chess that was discovered by the computer program designed for this research can be seen in Figure 2 where the bishop has just moved from d5 to e6. Since “mating squares” are also considered legitimate items that can be forked, this move qualified by threatening occupation of the f5 square and also the rook at h3.

It is not a typical fork since only one line is involved and the rook is attacked through the mating square but the threat is similar. There was nothing in the literature surveyed to exclude this type of position from being perceived as a fork so it was not invalidated. Such a fork however, will by default have two prongs. The only thing that might compromise the detection algorithm is multiple mating square threats along the same line. This would score unnecessarily high aesthetically for positions where checkmate could be delivered on say, any three adjacent squares on a line so mating square threats were limited to just one square per line.

3.1.9 The Pin/Skewer

A pin is in effect when a long range piece (i.e. bishop, rook or queen) attacks an enemy piece and prevents it from moving lest the more valuable or undefended piece behind it be captured. The main factors identified that differentiate one pin from another include the values of the pieces, distances between them and mobility of the pinned piece. The skewer is like an inverse pin. The more valuable piece is the one immediately attacked or “in front”. If both enemy pieces have the same value, it is still a skewer, not a pin. To ensure skewers and pins are mutually exclusive, pins are restricted to where the target piece (i.e. the one “behind” the pinned one) is worth more than the pinned one. The themes are evaluated as follows.

$$T_{2,3} = \frac{1}{p_c} \cdot \left[\begin{array}{c} \left(v\left(p_p\right) + v\left(p_t\right) + \frac{d\left(p_n, p_t\right)}{r\left(p_n\right)} \right) \\ - \left(\frac{m\left(p_p\right)}{r\left(p_p\right)} + k + l_a \right) \end{array} \right],$$

$$l_a = \begin{cases} \sum i\left(p_n, p_p\right) + \frac{v\left(p_n\right)}{\sum_1^n v\left(i\left(p_n, p_p\right)\right)}, \sum i \geq 1 \\ 0, \sum i = 0 \end{cases}$$

p_p = pinned/skewered piece, p_t = target piece, p_n = pinning/skewering piece, k = number of checking moves by p_p and p_t, l_a = (additional) liabilities, p_c = pin/skewer

constant (i.e. 19)
(9)

3.1.10 X-Ray

The x-ray theme occurs when an opponent's long-range piece is between two friendly long-range pieces (capable of defending each other) and can capture either one. It is more of a defensive maneuver than an attacking one since the x-rayed piece would have had to have been under attack by at least one of the x-raying pieces in the move prior. It is evaluated as follows.

$$T_4 = \frac{1}{x_c} \cdot \left(1 + \left| v(x_r) - \left(\overline{v(xp_1) + v(xp_2)} \right) \right| + \frac{d(xp_1, xp_2)}{\min\{r(xp_1), r(xp_2)\}} \right) \quad (10)$$

x_r = x-rayed piece, xp_1 = x-raying piece 1, xp_2 = x-raying piece 2, x_c = x-ray constant (i.e. 7)

The x-ray constant (to the nearest integer) is derived from an ideal x-ray of two queens x-raying a bishop across the board. This scenario, however unlikely, is nevertheless possible in chess and exhibits paradoxical features that are valued aesthetically. It is paradoxical because instead of removing the threat on the queen, another one is put *en prise* to create an x-ray. The inverse might also be seen as paradoxical in a sense (i.e. two bishops x-raying a queen) but this places more of an advantage to white (which we already know wins in the combination). It is not paradoxical in the right context because the victory becomes less of a surprise with white already having the advantage.

3.1.11 The Discovered/Double Attack

The discovered attack is a powerful tactic in chess where moving a piece uncovers an attack on an enemy piece. The discovered attack becomes a double attack if the moving piece uses the opportunity to attack another piece or the same one that is facing the discovered attack. If a double attack involves three or more pieces (e.g. a knight moves to create a discovered attack and simultaneously delivers a fork on two other pieces), only the more powerful of the two counts along with the discovered one. The fork will nonetheless register as a theme on its own. Any piece is capable of uncovering an attack on an enemy one so long as there is a long range piece behind it. It is evaluated as follows.

$$T_5 = \frac{1}{b_c} \cdot \left[\left(v(ba_m) + v(ba_k) + \frac{d(b_m, ba_m)}{r(b_m)} + \frac{d(b_k, ba_k)}{r(b_k)} \right) - (k + l_a) \right],$$

$$l_a = \begin{cases} \left[\sum i(b_k, ba_k) + \dfrac{v(b_k)}{\sum v(i(b_k, ba_k))} \right], \sum i \geq 1, v(ba_m) = 0 \\ 0, \sum i = 0 \end{cases}$$

ba_m = piece attacked by the moving piece, ba_k = piece attacked by the discovering piece, b_m = moving piece, ba_k = discovering piece, *i()* = intervening pieces
(11)

In the case of a double attack the moving piece is just as much a part of the theme as the one it uncovered (the "discovering" piece) so the main factor that aesthetically differentiates one instance of this theme from another is the combined value of the enemy pieces under attack. The theme constant, b_c is derived from an ideal instance of this theme, namely the double check (twice the value of the king).

3.1.12 The Zugzwang

This theme refers to positions where any move puts the player at a greater disadvantage than if he did not have to make a move. It usually occurs in the endgame where there are fewer pieces on the board and therefore also fewer legal moves available. Positions where it would disadvantageous for either player to move are called mutual or reciprocal zugzwangs. These are essentially still zugzwangs but from the other player's perspective. The main factor that differentiates one zugzwang from another is how intricate the position is or more accurately, how many possible variations or moves (all disadvantageous) are available to the player whose turn it is. A disadvantage here could mean checkmate, significant loss of material or a bad position where losing is ultimately inevitable.

$$T_6 = \frac{\sum z_m}{z_c}$$

z_m = (legal) moves available to the player in zugzwang
z_c = zugzwang constant (i.e. 30)
(12)

The formalization proposed (equation 12) is simply the number of possible moves available to the player in zugzwang, divided by the average number of moves in a typical chess position [4]. Even if it were known, the maximum number of moves possible in a legal chess position is not suitable because zugzwangs are rather limited to positions where moves are below average. The aesthetic value of a zugzwang therefore correlates with its complexity and improbability of occurring. Programmatically, it can be determined by permitting a second move to white (or null move by black) and looking for a forced mate-in-1. If there is none, it passes the test.

3.1.13 The Smothered Mate

The self-block or smothered theme involves checkmating the king with all of its flight squares blocked by friendly pieces, defended enemy pieces or a combination of both. Other major pieces (e.g. queen, rook) can also be smothered but this occurs less often because they are seldom worth the endeavour. The smothered mate can happen at any point in the game and is far more common in the corner and edge of the board than the centre because there are fewer flight squares. Due to its peculiar movement, the knight is often the checkmating piece in this theme but even a long range piece or pawn would suffice so long as it is defended against capture by the king. The proposed formalization is shown below.

$$T_7 = \frac{\sum r\left(s_p\right)}{s_c}$$

s_p = smothering pieces (those around the enemy king)
s_c = the smothered constant (i.e. 101)
(13)

The constant is derived from an ideal smothered mate at the center of the board with the maximum of eight pieces around the king. Based on the original piece set, the piece power of the most powerful 8 pieces, in order, are the queen (27), 2 rooks (14+14), 2 bishops (13+13), 2 knights (8+8) and a pawn (4). Since an occupied square is considered blocked regardless of piece type and it is precisely this blockage that is the main feature of the theme, only the number of pieces in the king's field count, not their value or colour. Consequently, smothered mates in the centre of the board score higher while those at the edge or corner score less.

3.1.14 Crosscheck

The crosscheck occurs when a player responds to a check with a reciprocal check. It is one of the few common chess themes where maneuvers by both players are taken into account for aesthetic purposes. The crosscheck is achieved by moving the king out of harm's way to uncover a discovered check on the opponent's king or intervening with a piece that simultaneously gives check. It does not usually include a check that results from capturing the checking piece. This rules out common positions where a series of checks is merely the result of repeated exchanges on the same square. Equation 7 shows how the theme score is calculated.

$$T_8 = \frac{c_n}{c_c}, c_c = \left[\left(2m\right) - 1\right], \quad m, n \geq 2$$

c_n = number of consecutive checks in the combination
c_c = crosscheck constant
m = number of moves in the combination
(14)

3.1.15 Promotion

Pawn promotion occurs when a pawn reaches the end of the board and promotes to either a queen, rook, bishop or knight. The most common choice is the queen but promotion to a knight is not uncommon, especially when it is prudent to do so. There are even cases where promotion to a bishop is necessary (e.g. where promoting to a queen gives stalemate) and promotion to a rook results in a faster win. One of the best examples of the latter is the Saavedra position from the late 19th century. Underpromotion is considered more beautiful because it is both paradoxical and economical. The formalization is therefore given as:

$$T_9 = \frac{p_c}{v\left(p_p\right)}$$

p_c = promotion constant (i.e. 3)
p_p = promotion piece
(15)

3.1.16 Switchback

The switchback is the return of a single piece to its initial square (either immediately or later in the move sequence). For the purpose of this research, it also includes the similar *rundlauf* or round-trip theme where a piece leaves a square, and then later in the solution returns to it by a circuitous route (e.g. a rook moves e3-g3-g5-e5-e3) whereas in the switchback, a piece leaves a square and then later in the solution returns to it by the same route (e.g. a rook moves e3-e5-e3). Only pawns are incapable of such a maneuver. Given the scope, this theme can only occur once in the move sequence.

Distance and piece power are the main aesthetic factors but they only play a role if the "all of the pieces power" (P_3) theme is not evaluated to avoid redundancy. In that case the score is the total distance traversed by the switchback piece, to each of the squares in its path, over its power. This is essentially the same as P_3.

3.2 Model of Aesthetics

The formalizations for the principles just described are not enough on their own to evaluate aesthetics in chess, even mate-in-3 combinations even though they might be capable of identifying highlights of a particular move sequence. A cumulative model of aesthetics is therefore proposed in the form given below.

$$A = \sum_1^m P_m + \sum_1^n T_n$$

A = aesthetic value of a combination,
P = aesthetic principle evaluation score,
T = theme evaluation score

(16)

The sum of aesthetic principles and themes present in a combination should theoretically be higher for beautiful ones. It stands to reason that attractive or 'brilliant' move sequences in real games and compositions should contain not only more aesthetic principles and themes but better instances or configurations of them which the formalizations proposed are flexible enough to evaluate. The presence of more however, does not guarantee a high score (their individual evaluations may be low) and neither does few guarantee a low score (individual evaluations may be high).

4 Experimental Results

A computer program called, CHESTHETICA was developed incorporating the aesthetics model because manual evaluation is tedious and prone to error. The program does not possess any game playing intelligence but is capable of facilitating a match (with all the special rules e.g. castling, *en passant*, promotion) between two players. This was necessary to set the foundation for proper evaluation of all the aesthetic principles and detection of relevant themes.

Several novel experiments were designed to see if a computer program incorporating the model would generally rate chess compositions higher than regular games in terms of aesthetics, consistent with human perception of beauty in chess. For this purpose, 4 sets of randomly selected mate-in-3 chess compositions and similar combinations from actual tournament games were used. Both compositions and real games each consisted of two sets of 3,000 combinations (for a total of 12,000). Because aesthetics in chess tends to correlate with sound play, only games between expert players (ELO rating ≥ 2000) were used. The ELO rating system is a widely used method for calculating the relative skill of chess players.

Novice and intermediate play would inherently be less beautiful and bias the results. It is true that most master games end with one player resigning (as opposed to being checkmated) but given the wide availability of games in commercial databases, a sufficient number could be found for the experiments. A resigned game with an inevitable forced mate could also have been used with the aid of a computer to find the mating variation. Most resigned master games however, are not that close to checkmate. The expert game checkmates used in the experiments were not necessarily forced mates like the compositions because this does not affect its aesthetic evaluation in any way. The important thing is that the checkmates were played out in full by humans and not generated artificially by a computer. Table 1 and 2 show the results obtained. They have been sorted in descending order for clarity.

Real Games		**Compositions**	
Set 1	**Set 2**	**Set 1**	**Set 2**
1.802 SD 0.711	1.800 SD 0.716	2.665 SD 0.871	2.689 SD 0.873

Table 1: Mean aesthetic scores for sets of 3,000 combinations

RG1 vs RG2	**CP1 vs CP2**	**RG1 vs CP1**	**RG2 vs CP2**
Not Significant	Not Significant	t(5766) = -42.1, P<0.001	t(5777) = -43.2, P<0.001

Table 2: Significance of differences between mean aesthetic scores

Table 1 shows the mean aesthetic scores between the collections of real game combinations and compositions. The difference between real games sets is negligible as it is between compositions. This was expected. Between compositions and real games however, both sets demonstrated a stark difference in aesthetic scores, consistent with expectations that compositions are generally more beautiful. The standard deviations of the scores are relatively high, indicating distinct variation in the combinations, consistent with aesthetic evaluation, especially in compositions. Table 2 shows that the minute difference between the two sets of real games and difference between the sets of compositions was not significant (two-sample t-test assuming unequal variances). However, the stark differences between each set of real games vs compositions was significant.

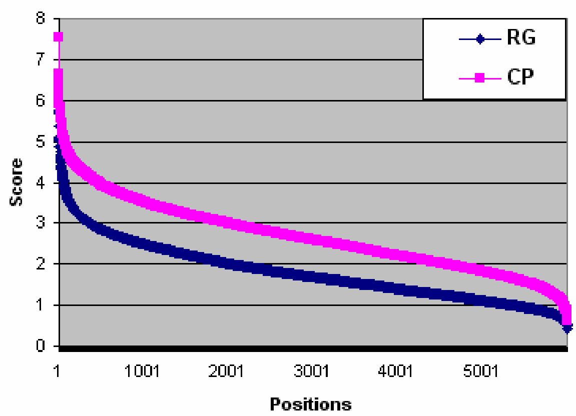

Figure 3

Figure 3 shows the distribution of scores arranged in descending order for illustrative purposes. The implications of these results and possible applications of this research are discussed next.

5 Discussion

The statistically significant differences in means found between the aesthetic values of chess compositions and real games suggest that aesthetics in chess can be recognized computationally. This does not mean that compositions necessarily score higher than real games in terms of beauty because there are always exceptions such as poorly composed problems and overrated combinations in real games. Figure 3 clearly shows that there are combinations from real games that score higher than some compositions. Even so, a high score based on the aesthetics model proposed would likely point to a move sequence that humans would find beautiful.

Experiments involving human players were not performed because their knowledge of what constitutes beauty in chess would be difficult to ascertain as reliable (e.g. like that found in chess literature). It is difficult to determine if shorter or longer move sequences would exhibit similar aesthetic scores because shorter ones tend to be quite simple (and limited thematically) whereas longer ones can get rather complicated and difficult for humans to follow. Comparisons between move sequences of different lengths are not as reliable for the same reason. Modifications or extensions to the model could be applied to compensate for these possible discrepancies.

A comparison against the traditional approach of attributing fixed values to themes and principles was not done because the selection used here are not adequately represented in prior work which focus mainly on compositions and their conventions [19][20[23]. Since the evaluation functions proposed are scalable and can cater for many different configurations of individual themes and principles, they are nevertheless assumed to be better. This is supported by the results of the experiments performed.

Chess database search engines can incorporate the aesthetics model proposed to locate aesthetically pleasing combinations in vast databases of games for human appreciation and study. Automatic problem composers can also use the formalizations presented to improve their fixed-value approach to aesthetics and to decide, without human intervention, which derived forced checkmates are the best. In addition, chess composition and brilliancy prize judges might find some impartial assistance through this model when deciding on a winner [44]. Finally, complex compositions could more quickly be solved (and sometimes solved *at all*) if chess game tree search heuristics employed heuristics based on aesthetics because brute-force and common pruning techniques often overlook paradoxical but necessary key moves [23].

5.1 Aesthetics in Other Games

Investigations into the game of chess have often unintentionally yielded benefits in other domains but the game itself should not be seen as nothing more than a stepping stone toward greater things. With millions of players worldwide and constant efforts to improve computer playing ability, this research opens another facet of inquiry (more commonly known as computational aesthetics) that has its own benefits, particularly to chess players and composers.

It shows that there is still much to learn about what humans like and experience in the game. This can be potentially enhanced through the use of computers in the same way that rigorous computer analysis has revolutionized many of our old ideas about how the game should be played. Unlike checkers, chess is still far away from being solved [45]. AI researchers have nothing to be ashamed about if their research into chess translates into technologies that mainly have the potential of enriching the experience of human players.

Nevertheless, other sufficiently complex (i.e. having an aesthetic component or domain) zero-sum perfect information games such as go can apply the same methodology used here to develop their own aesthetic models once there is enough literature to substantiate it [46]. A direct application of the formalizations presented here is not really possible (with the exception of *sparsity*) because aesthetics in such games is inextricably linked to the rules which govern them. For example, in (Western) chess there are 6 piece types whereas in go there is only one.

Therefore visual pattern recognition would most likely be more significant to the aesthetics of go than it is in chess. Economy on the other hand would be much less about piece values than it is about mobility given the objective of go which is to control more territory on the board. Similar to chess, the aesthetics of go is associated with sound play and could contribute to the development of better game playing technologies. When computers are able to beat the strongest human go players, attention might shift to aesthetics for its own sake like is being done now in chess [47][48].

Chess variants, estimated at over 1000 in number, would be more amenable aesthetically to the formalizations in this paper because only minor modifications would be required to adapt them [49]. Many variants were in fact created due to aesthetic limitations in the original game. For instance, variants that use fairy chess pieces (unorthodox ones not in the standard set) or different board types could easily derive their piece values and mobility according to the methods clearly described in section 3.

Finally, it is difficult to say if aesthetic recognition in board games could also contribute to the humanization of otherwise bland and brutal game playing software. Associating a kind of emotional response in programs that would favour say, making the beautiful move - even when it is not necessarily the most effective one - could bring us a step closer toward that objective. It is important to note that this is quite different from opponent “personalities” (available in certain programs) that essentially limit playing strength rather than take notice of what is aesthetically preferential, like humans often do [50].

6 Conclusions and Further Work

In this paper, formalizations for established aesthetic principles in chess were proposed and cumulatively presented as an aesthetics model for the game. This model was incorporated into a computer program and used to compare large samples of randomly selected direct mate-in-3 compositions and mate-in-3 combinations from expert-level tournament games. The results showed a statistically significant difference in their means suggesting that computers can use the model to recognize beauty in the game.

The aesthetics model can be further enhanced by including formalizations of additional aesthetic principles and individual formalizations for a wider variety of chess themes. Chess literature places no emphasis on particular aesthetic principles so not weighting them individually minimizes bias. The aesthetics model and formalizations within are flexible enough to cater for shorter and longer combinations but a new set of experiments would be required when comparing sequences of different lengths because they are perceived differently by humans [30].

Applications of this research are most obvious within the domain of chess but extensions to other games of similar complexity are entirely feasible. Very recently, the authors were contacted by another researcher (also an International Chess Master) who expressed keen interest in these aesthetic evaluations for the purpose of enhancing an automatic chess game commentator under development [51]. There is likely potential for further collaborative work in that respect, and possible enhancement of the model presented here. With sufficient processing power, it is quite possible that computers will one day be able to discover amazing and brilliant combinations in the game tree for human aesthetic appreciation and study that would otherwise take centuries to occur in real games or be thought of by composers.

7 Acknowledgements

I would like to thank John McCarthy (Stanford), Michael Negnevitsky (University of Tasmania), Jaap van den Herik (Universiteit Maastricht), John Troyer (University of Connecticut), Malcolm McDowell (British Chess Problem Society), Brian Stephenson (Meson Database), GM Jonathan Levitt and IM David Friedgood for their comments and feedback.

8 Appendix

Chess problems obtained from Meson Database (26558 #3 problems);
http://www.bstephen.me.uk/access_meson.html

FIDE Master tournament games obtained from ChessBase MegaDatabase 2008 (3803334 games);
http://www.chessbase.com/shop/

References:

[1] Kasparov, G., *Kasparov Teaches Chess*. B. T. Batsford, London, 1987.

[2] Newborn, M., *Kasparov versus Deep Blue*. Springer-Verlag, New York, 1997.

[3] Levy, D., *How Computers Play Chess*. W. H. Freeman & Co (Sd), 1991.

[4] Shannon, C. E., Programming a Computer for Playing Chess. *Philosophical Magazine*, Ser. 7, Vol. 41, No. 314, 1950, pp. 256-275.

[5] Polgar, S., Chess – a world phenomenon, *ChessBase News*, 19th April 2005. Available at: http://www.chessbase.com/newsdetail.asp?newsid=2336

[6] Lasker, E., *Lasker's Manual of Chess*. Dover Publications Inc., New York, 1947.

[7] Margulies, S., Principles of Beauty, *Psychological Reports*, 41, 1977, pp. 3-11.

[8] Lionnais, F. Le. *Les Prixde Beaute aux Echecs*. 2nd edition. Payot., 1951.

[9] Troyer, J.G. Truth and Beauty: The Aesthetics of Chess Problems. In *Haller (ed.), Aesthetics* (Holder-Pichler-Tempsky, Vienna): 126-30, 1983.

[10] Bronstein, D., *Chess in the Eighties*. Macmillan Pub Co., England, 1983.

[11 Levitt, J. and Friedgood D., *Secrets of Spectacular Chess*. Henry Holt and Company Inc., New York, 1995.

[12 Osborne, H., Notes on the Aesthetics of Chess and the Concept of Intellectual Beauty. *British Journal of Aesthetics* 4, 1964, pp. 160 - 163.

[13] Humble, P.N., Chess as an Art Form. *British Journal of Aesthetics*, 33, 1993, pp. 59-66.

[14] Ravilious, C. P., The Aesthetics of Chess and the Chess Problem. *British Journal of Aesthetics*, 34, 1994, pp. 285 - 290.

[15] Damsky, I., *Chess Brilliancy*. Everyman Publishers, London, England. 2002.

[16] Iqbal, A., Is Aesthetics Computable? *ICGA Journal*, 29(1), 2006, pp. 32-39.

[17] Schlosser, M., Computer and Chess Problem Composition. *ICCA Journal*, 11(4), 1988, pp. 151-155.

[18] Schlosser, M. Can a Computer Compose Chess Problems? *Advances in Computer Chess 6*, Ellis Horwood Ltd., Chichester, UK, pp. 117-131., 1991.

[19] HaCohen-Kerner, Y., Cohen, N. and Shasha, E., An Improver of Chess Problems. *Cybernetics and Systems*, 30(5), 1999, pp. 441-465

[20] Fainshtein, F. and HaCohen-Kerner, Y. A Chess Composer of Two-Move Mate Problems. *ICGA Journal*, 29(1), 2006, pp. 3-23.

[21] Howard, K. *The Enjoyment of Chess Problems*. Dover Publications Inc., New York, 1967.

[22] Iqbal, A. Computing the Aesthetics of Chess. Technical Report of the *AAAI Conference/(M.I.T) Workshop on Computational Aesthetics*, Boston, AAAI Press, USA, 16-20 July, 2006, pp. 16-23.

[23] Walls, B. *Beautiful Mates: Applying Principles of Beauty to Computer Chess Heuristics*. M. Sc. Dissertation. University of Sussex, UK, 1997.

[24] Gomboc, D., Buro, M. and Marsland, T., Tuning Evaluation Functions by Maximizing Concordance. *Theoretical Computer Science*, 349 (2), 2005, pp. 202-229.

[25] Datta, R., Joshi, D., Li, J. and Wang, J., Studying Aesthetics in Photographic Images Using a Computational Approach. *Proceedings of the European Conference on Computer Vision*, Part III, Graz, Austria, May 7-13, 2006, pp. 288-301.

[26] Manaris, B., McCormick, C. and Purewal, T. Can Beautiful Music be Recognized by Computers? *Technical Report CoC/CS* TR#2002-7-1, March 1st., 2002.

[27] Manaris, B., McCormick, C. and Purewal, T. 2002. Progress Towards Recognizing and Classifying Beautiful Music with Computers by Bill Manaris; *IEEE SoutheastCon Conference*, Columbia SC , ETATS-UNIS 4th May, 2002, pp. 52-57

[28] Euwe, M. *The Logical Approach to Chess*. Dover Publications Inc., New York., 1982.

[29] McDowell, M. Selfmates. *British Chess Problem Society*, 2007. Available at: http://www.bcps.knightsfield.co.uk/selfmates.html

[30] Groot, A., *Thought and Choice in Chess*. Moulton. New York., 1965

[31] Akl, S.G. and Newborn, M. The Principal Continuation and the Killer Heuristic. *ACM*

Annual Conference Proceedings, ACM, Seattle, USA. 1977, pp. 466–473.

[32] Hartmann, D., How to extract relevant knowledge from grandmaster games, Part 1: grandmasters have insights—the problem is what to incorporate into practical programs, *ICCA Journal*, 10 (1), 1987, pp. 14–36

[33] Hartmann, D., How to extract relevant knowledge from grandmaster games, Part 2: the notion of mobility, and the work of De Groot and Slater, *ICCA Journal*, 10 (2), 1987, pp. 78–90.

[34] Hartmann, D. Notions of evaluation functions tested against grandmaster games, in: D.F. Beal (Ed.), *Advances in Computer Chess*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1989, pp. 91–141.

[35] Harding, T. Materialism in Chess. *The Kibitzer*, September, 2002. Available at: Chess café. http://www.chesscafe.com

[36] Gallagher, J. *The Magic of Mikhail Tal*. Everyman Chess, 2001.

[37] Nunn, J. *Solving in Style*. Gambit Publications Ltd., USA, 2002.

[38] Zirkwitz. T. A Short Introduction to the Bohemian Chess Problem. *The Problemist Supplement*, Issue 10, 1994, pp. 78-80.

[39] Iqbal, A. Evaluation of Economy in a Zero-sum Perfect Information Game, *The Computer Journal*, doi: 10.1093/comjnl/bxm060, 2007

[40] Woodruff, A. VIDA (Visual Information Density Adjuster). *CHI 99* 15-20, ACM, May. 1999.

[41] Grand, H. The Impact of Computers on Chess Problem Composition. *ICCA Journal*. September, 1986, pp. 152-153.

[42] Silman, J. *How to Reassess Your Chess*. Siles Press. USA, 1993.

[43] Rice, J. Chess Wizardry: *The New ABC of Chess Problems*. International Chess Enterprises, USA, 1997.

[44] Albrecht, H. How Should the Role of a (Chess) Tourney Judge Be Interpreted? *The Problemist*, Vol. 7, 2000, pp. 217-218.

[45] Schaeffer, J., Burch, N., Björnsson, Y., Kishimoto, A., Müller, M., Lake, R., Lu, P., and Sutphen, S. Checkers is Solved, *Science* 14 September, 317: 2007, pp. 1518-1522

[46] Guosun, S. *Beauty and the Beast: Exquisite Play and Go Theory*. Yutopian Enterprises, Santa Monica, CA., 1997.

[47] Johnson, G. To Test a Powerful Computer, Play an Ancient Game. *New York Times*: Cybertimes. July 29, 1997.

[48]Hsu, Feng-Hsiung, Cracking Go, *IEEE Spectrum*, Volume 44, Number 10, USA, 2007

[49] Pritchard, D. *Popular Chess Variants*. Batsford Chess Books, London, 2000.

[50] Chessmaster 11. *Ubisoft*. 2007. Available at: http://chessmaster.uk.ubi.com/xi/index.php

[51] Guid, M., Mozina, M., Krivec, J., Sadikov, A. and Bratko, I. Learning Positional Features for Annotating Chess Games – A Case Study. *Proceedings of the Conference on Computers and Games 2008*, Beijing, China, 29 September to 1 October 2008 (accepted paper)