Increasing Efficiency and Quality in the Automatic Composition of Three-Move Mate Problems

Azlan Iqbal

College of Information Technology, Universiti Tenaga Nasional, Kampus Putrajaya,
Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia
azlan@uniten.edu.my

Abstract. In this article, we demonstrate the use of composing 'experience' in the form of piece location probability values derived from a database of mate-in-3 chess problems. This approach was compared against a 'random' one. Comparisons were made using 'experiences' derived from three different databases, i.e. problems by human composers (HC), computer-generated compositions that used the HC experience (CG), and mating 'combinations' taken from tournament games between humans (TG). Each showed a reasonable and statistically significant increase in efficiency compared to the random one but not each other. Aesthetically, the HC and CG were better than the others. The results suggest that composing efficiency and quality can be improved using simple probability information derived from human compositions, and unexpectedly even from the computer-generated compositions that result. Additionally, these improvements come at a very low computational cost. They can be used to further aid and entertain human players and composers.

Keywords: Artificial intelligence, chess, composition, probability, experience, efficiency, aesthetics.

1 Introduction

A chess problem or 'composition' is a type of puzzle typically created by a human composer using a chess set. It presents potential solvers with a stipulation, e.g. *White to play and mate in 3 moves*, and is usually composed with aesthetics or beauty in mind. Compositions often adhere to certain composition conventions as well, e.g. *no 'check' in the key (i.e. first) move*. One of the earliest books on chess problems, using an early Indian form of the game, is from the 9th century AD [1]. Composition tournaments are at present held all over the world and attract competitors from diverse backgrounds [2].

The automatic composition of chess problems – pertaining to Western or international chess, in particular – is relatively uninvestigated. Especially in contrast to chess *playing* which was "*once seen as a really hard thing humans could do and computers couldn't*" [3] but is now losing emphasis in the artificial intelligence (AI) community in favor of 'more complex' games like go and Arimaa [4-6]. Perhaps a better reason than the emphasis on chess playing for much of AI history, automatic

chess problem composition may have suffered because 'creativity', the essential component that is usually mirrored in aesthetics, is not well defined [7].

Chess problems actually provide a convenient domain of investigation for creativity or aesthetics – more so than say, music [8] or paintings [9] – since there is a clear and distinctly measureable contrast between problems by human composers and otherwise-identical move sequences or 'combinations' that typically take place in real games between humans [10, 11]. Advances in this area can be of immediate benefit to human composers and players in terms of educational and entertainment value since there is virtually no limit to the potential output of a machine.

Section 2 reviews briefly related and relevant previous work. Section 3 details the steps involved in the composing approaches tested. Section 4 explains the experimental setup and results. Section 5 presents a discussion of the results. Section 6 summarizes the main points with some thoughts on future work.

2   Previous Work

Schlosser presented an effective 'basic' method of automatically composing chess problems [12, 13]. It consists essentially of: 1) constructing a complete database of chess endgames, 2) eliminating positions that do not have a unique and optimal move sequence, and 3) selecting the 'true' chess problems with the help of a human expert. The method is therefore limited to the number of pieces which endgame databases or 'tablebases' support, e.g. presently 6 pieces including kings [14], and also depends on human expert intervention. This inherently limits the 'automatically' generated compositions in terms of scope and output potential.

Noshita explained how – for Tsume-Shogi (Japanese chess mating problems) – a board position can be randomly-generated and then 'transformed' by removing or adding pieces through certain operations [15]; a principle equally applicable to international chess. A game-solving engine can then be used to determine the solution, if any. The majority of such positions end up having no solution. Some improvements, e.g. generating compositions with more pieces, can be obtained by 'reversing' moves one at a time from a given mate position with the aid of a solving engine to test if the problem is 'complete', i.e. it has a unique solution [16].

Several criteria for determining the artistic value of the compositions can be automatically tested for but these tend to rely on confirmation by human experts, in any case [16]. The reason is likely because these criteria and their automatic evaluation techniques may not have been experimentally-validated. 'Optimization' or the removal of unnecessary pieces can be performed so long as it does not invalidate the solution. Pieces may also be added at particular points in the process [17]. Within a restricted scope, a 'reverse' method without the need for a time-consuming solving engine is possible [18], but this typically comes at the cost of more memory.

Chess problems, e.g. two-movers, can also be 'improved' a fair amount using a computer. Domain experts are first consulted in order to formalize the knowledge required to assess the 'quality' of compositions. This results in simple formulas (e.g. *first move pins a white piece = 3 × piece's value*) or specific weights for certain detectable maneuvers and features (e.g. *Grimshaw theme = 45*). Pieces can then be deleted, added or replaced to improve the position [19]. Composing efficiency and

quality here can be improved but mainly via somewhat computationally-intensive 'search' enhancements [20, 21].

In summary, the two main issues in this area relate to computational efficiency (processing power, memory) and quality functions (aesthetics). The first is reminiscent of the problem the AI community faced with regard to early chess playing programs, but eventually 'solved' thanks in large part to powerful hardware that became available. Former world chess champion Garry Kasparov's loss to IBM's supercomputer Deep Blue in 1997 being the prime example. That, in retrospect, is not seen as much of an achievement in AI. It is therefore preferable not to fall into the same predicament with compositions. The output of automatic composers, relative to the time they take, should be increased at minimal computational cost; this can be seen as making them 'cleverer'.

The second issue is likely due to the lack of experimental validation when it comes to quality functions and aesthetics models, and an over-reliance on expert opinion which tends to be inconsistent [11, 22, 23]. There is likely no consistent 'methodology' pertaining to how the best human composers compose problems. They tend to take their time, abide by a number of 'accepted' composition conventions, and leave the rest to personal style, experience and creativity [24]. This may be why it is difficult for computers to compose original chess problems like they do, and to do so on demand or within a stipulated time frame. In any case, aesthetics or quality in chess can, in fact, now be *assessed* computationally to a reasonable degree within the scope of three-move mate problems using an experimentally-validated model [10, 11, 25]. This minimizes or removes altogether the need for human expert intervention for chess problems of that type, and makes the process of aesthetic assessment more consistent, reliable and affordable; especially for research purposes. That model will therefore be used in lieu of chess experts, perhaps for the first time, to assess the quality of our automatically generated compositions. More details about it are in section 4. The ability to generate more efficiently compositions that 'work' is good in itself but that they are, on average, of higher quality is even better.

3 The Composing Methodology

This research was limited to orthodox mate-in-3 problems (#3) in standard international chess. The 'composing' feature was incorporated into a computer program. Two automatic composing approaches were compared, i.e. 'experience' and 'random'. The main difference between them is that the first uses an 'experience table', explained below, after step 4. The process as follows applies to both.

1. Place the two kings on random squares on the board. Accept them so far as the resulting position is legal; otherwise, repeat the process.
2. Alternating between White (first) and then Black, determine whether the next selection will be an actual piece or a 'blank', i.e. nothing. The probability of choosing a blank for White was set to 16.67% (1 in 6 chance, given the other five piece types) whereas for Black, it was set to 33.34% (twice as likely) to give White a better chance of checkmating ; compositions are generally seen from the standpoint of White winning.
3. If a 'blank', return to step 2 with the side in question having lost its 'turn'.

4. If an actual piece is to be selected, choose one of the five remaining piece types at random (equal odds) and place it on a random square that is unoccupied. Keep trying until one is found.

This is where the two approaches diverge. In the random approach, no position transformation occurs and we skip to just after step 7. In the experience approach, the experience table is used for that purpose. The table is created (beforehand) based on the rapid, automatic analysis of a chess problem database. Three databases (see section 4 for details) were used in experimentation to derive three different experience tables. Fig. 1 shows, in two columns, how the probability information may be stored in a text file.

$$Sq: \ 36 \qquad 3: \ 1.78$$
$$0: \ 65.33 \qquad 4: \ 1.72$$
$$1: \ 4.8 \qquad 5: \ 0.66$$
$$2: \ 4.33 \qquad 6: \ 0.52...$$

Fig. 1. Contents of an 'experience table'

The first line indicates the square (0-63) – upper left to lower right of the board – followed by the piece types (0-12) and their associated probabilities (% occurrence) as follows: blank, (white) pawn, knight, bishop, rook, queen, king, (black) pawn, knight etc. For instance, in this example it was determined that, in the initial positions, a white bishop occupied one of the central squares (*e4*, in the algebraic chess notation) only 1.78% of the time.

5. Based on the probability information, examine the squares immediately around the one chosen in step 4 for potentially better placement.
6. If there is a king on one of those squares, skip it. If a square is blank but has a higher associated probability value for the random piece selected in step 4, shift the random piece there.
7. If there is a piece in a surrounding square but that square has a higher associated probability value for the random piece than the one currently on it, replace it with the random one. In order to increase the likelihood of White being able to force mate, black pieces cannot replace white ones.

At this point, the two approaches converge. The 'final' generated position was set to have at least two black pieces to avoid 'lone king' mates, a minimum total of four pieces and a maximum total of sixteen pieces.

8. If a minimum piece requirement is not satisfied, return to step 2.
9. If the maximum piece limit is exceeded, discard the position thus far and return to step 1.

The following are some possible 'violations' in the composing process.

a. Exceeding the original piece set, e.g. having three rooks of the same color.
b. Having two bishops in an army occupying squares of the same color.
c. Having a pawn occupying the eighth rank.

The first two are not actually illegal but they are somewhat unconventional in compositions. In such cases, the 'offending' piece is removed and the process returns to step 4; this therefore does not count as a further transformation of the position but simply a 'mistake' to be corrected. The possibility of castling was given a 'neutral' 50% random probability of being legal, assuming a king and one of its rooks happen to be on the right squares. Determination of legality based on retrograde analysis was considered unnecessary for the purposes of this research [26, 27]. En passant captures, if plausible, default to illegal. 'Officially', in compositions, castling in the key move is legal unless it can be proved otherwise whereas en passant is legal only if it can be proved the last move by the opponent permitted it [28].

10. If an illegal position results, remove the random piece from its square and return to step 2.
11. A mate-solver is used to determine if the tentatively acceptable position generated has a forced mate-in-3 solution to it. If not, do not remove the random piece, and return to step 2.
12. If there is such a solution, the position is optimized as shown in the code below. This makes the composition more economical in form [29].

```
FOR every square
    IF not occupied by a king and not empty THEN
        Remove piece
        IF forced mate-in-3 can still be found THEN
            Proceed
        ELSE
            Return piece to its original location
        END IF
    END IF
NEXT
```

To be thorough, optimization is performed three times, starting from the upper left to the lower right of the board; white pieces first, then black, and then white again. Fewer passes proved to be insufficient in certain positions. Optimization generally increases the aesthetic quality of a composition by removing unnecessary or passive pieces and should apply equally to both the random and experience approaches to make the comparisons more meaningful.

13. If the position can be optimized, test it as in step 8. Satisfying that, consider the composing attempt successful.

The number of transformations or iterations per composing attempt was limited to 21, after which a new composing attempt begins regardless (step 1). Scant positions would result given too few iterations, and the opposite given too many. There was no implementation of particular composition conventions, e.g. *no captures in the key move*, *no 'duals'* (see section 5 for more on this). The process as described in this section may seem more complicated than necessary. For instance, why not just draw the pieces from a fixed set and place them on the squares based on their associated probabilities in the experience table? The reason is that doing so results in less creative variation and very similar-looking, if not identical, generated compositions.

4   The Experimental Setup

For every comparison made, automatic composition was attempted 100 times for 40 'cycles'. The composing efficiency (i.e. successes/attempts) for each cycle was calculated, and the mean used as a basis of comparison. For comparisons within the experience approach, i.e. between different experience tables, the databases included: 29,453 (mostly published) problems by human composers[1] (HC), 1,500 computer-generated compositions that used the experience table derived from the HC (CG), and 3,688 forced mate-in-3 combinations taken from tournament games between at least club-level human players, i.e. rated at least 1600 Elo points (TG). From this point, mention of any of these databases in the context of composing efficiency will refer to the experience table that was derived from it. For statistical purposes, usage of the two sample t-test assuming equal (TTEV) or unequal (TTUV) variances – to establish if a difference in means was significant – was determined by first running a two sample F-test for variances on the samples (which were assumed to have a normal distribution). T-tests were all two-tailed, and at a significance level of 5%. Table 1 shows the results. The standard deviation is given in brackets.

Table 1. Mean composing efficiency.

Random	Experience		
	HC	CG	TG
23.70%	28.03%	28.38%	27.25%
(3.69)	(4.45)	(4.49)	(3.94)

The differences in mean efficiency were *not* statistically significant between any of the experience approaches. However, they were all different to a statistically significant degree when compared to the random approach as follows.

HC vs. Random: TTEV; $t(78) = 4.735$, $P<0.01$
CG vs. Random: TTEV; $t(78) = 5.087$, $P<0.01$
TG vs. Random: TTEV; $t(78) = 4.160$, $P<0.01$

Even though the improvements may not *look* very large in terms of raw percentage, they actually translate to quite a few more successful compositions than the random approach. For instance, after 10,000 composing attempts, the CG approach would have 433 *more* compositions than the random one. This is enough to fill two small books on chess problems.

Every automatically generated composition was assessed using the chess aesthetics program, CHESTHETICA that incorporates Iqbal's mate-in-3 chess aesthetics model [10, 11, 25]. The model is too complex to be sufficiently explained here but all the necessary information pertaining to its workings are available in the resources just cited. We are not attempting to further debate or justify its merits here but simply consider it validated and are applying it to *this* research. In principle, the model uses

---

[1] Sourced from Meson Chess Problem Database (http://www.bstephen.me.uk/); courtesy of Brian Stephenson.

17 aesthetic features (e.g. *pin, skewer, fork, economy, sparsity, material sacrifice*) common to real games and compositions, and can discriminate effectively *between* these domains but not *within* them. To give the reader some idea how the features are evaluated, one of the evaluation functions is shown in equation 1 where T_1 represents the aesthetic value of the *fork* theme, if detected, after a piece moves; $v()$ denotes the standard Shannon value of the piece, $d()$ the Chebyshev distance between two pieces and $r()$ the 'power' of the piece. These concepts are fully explained in [10, 11, 25].

$$T_1 = f_c \times [(\sum v(fp_n) + n + (\sum d(f_k, fp_n) \times r(f_k)^{-1})) - k] . \tag{1}$$

f_c = fork constant, fp = forked piece, f_k = forking piece,
k = number of possible 'check' moves by fp

The results also correlate well with mean human-player aesthetic ratings and agree with the typical selections of experts. It is a more consistent, reliable and cost-effective alternative to the traditional approach of using one or two human experts. This is something that was not possible before due to a lack of experimentally-validated aesthetic assessment technology. Given the sheer number of compositions to be assessed, human experts would not have been a viable option here in any case. A higher score implies that the combination is more likely to be considered beautiful by the majority of human chess players of reasonable competence in the game. The model is used to evaluate *beauty* in the game, and therefore does not explicitly account for some of the composition conventions – that may have little to do with beauty or creativity per se – typically adhered to by composers. Table 2 shows the mean aesthetic scores. The standard deviation is given in brackets.

Table 2. Mean composition aesthetic scores.

Random	Experience		
	HC	CG	TG
2.104	2.168	2.178	2.088
(0.44)	(0.46)	(0.46)	(0.46)

The differences in means were *not* statistically significant between the HC and CG approaches, and TG and random, but were in all other cases as follows.

HC vs. Random: TTEV; $t(2067) = 3.259$, $P<0.01$
CG vs. Random: TTEV; $t(2081) = 3.743$, $P<0.01$
HC vs. TG: TTEV; $t(2209) = 4.121$, $P<0.01$
CG vs. TG: TTEV; $t(2223) = 4.619$, $P<0.01$

Even though the difference between say, the CG and random approaches is small, i.e. 0.074 – and therefore probably difficult for humans to perceive – it is nevertheless an improvement over the random approach, and no worse than it despite the reasonable increase in efficiency (see Table 1). An aesthetic difference of approximately 0.5 or more would be more obvious to humans [25]. Mate-in-3 compositions by humans average approximately 2.1 aesthetically whereas analogous combinations from tournament games average about 1.7 [25]. The latter, in their 'original' form, are not optimized in any way,

e.g. step 12 in section 3. The aesthetic score for a combination from either of these domains typically ranges from 0.5 to 5.0; in rare cases slightly lower or higher. Overall, these results suggest that the experience approach, using a table of piece-placement probability values derived from any of the databases is reasonably better than the random approach in terms of composing efficiency, and to a small but significant degree also aesthetics (given the HC and CG).

5   Discussion

The experimental results show that simple piece-placement probability values derived from a database of compositions – when used in the composing process as explained in section 3 – improves composing efficiency compared to an approach that does not benefit from that information. In two of the three experience approaches (HC and CG), a small improvement in terms of aesthetics was also detectable. The CG approach was actually included to see if the computer could also 'learn' from its *own* composing experience but this does not appear to be the case. The computational cost for these improvements is minimal, especially in contrast to any approach that involves game-tree searching [30] beyond confirming if a solution exists, i.e. the use of a solving engine.

Previous work in the area (see section 2) may not have explored this idea to improve efficiency and aesthetics because sizeable databases of human compositions are generally difficult to come by. Without the appropriate sort of piece placement and position transformation process (see section 3), the approach may also have seemed likely to converge toward a 'local maximum', i.e. the high probability of certain pieces on certain squares and subsequently similar compositions generated. In any case, until the idea is actually tested as was done here, we cannot be sure what the results would look like. For instance, without experimentation it would have been difficult to predict that 'experience' gained from a TG database would improve composing efficiency (though not aesthetics), or that compositions generated by a computer based on experience derived from human compositions would be just as effective as a *source* of experience.

Even though the generated compositions benefited from the probability values derived from a database of problems by human composers (HC), that database henceforth becomes unnecessary given the (smaller but equally effective) computer-generated one it helped produce (CG). The reason for this may be because there is a lot of personal style, taste and convention in human compositions that do not necessarily aid the strict process of composing problems that 'work' or even relate to beauty per se [11]. This 'information' may be mostly stripped away in the automatically generated compositions that result. It is important to remember that a winning composition is not necessarily among the most *beautiful*. Sometimes, though not often, the elegance of a simpler composition or an analogous combination that occurs in an actual game can be considered more beautiful by the majority of reasonably competent players than a 'difficult' composition with hundreds, if not thousands, of variations that only the most experienced composers can appreciate. Fig. 2 shows an example of this contrast; one is a 1st prize winning composition by a human composer (rather complicated) and the other occurred in a game between two chess *engines* in a simulated match (simple and elegant).

1. Qf8 exd6+ 2. Qe7 Nd7 3. Qxd6#
Alfreds A Dombrovskis,
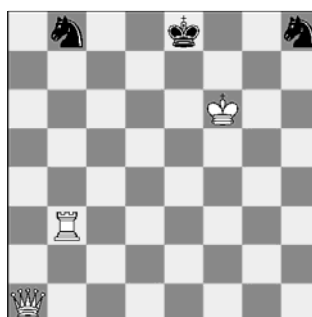1st Prize, Schakend Nederland
1973
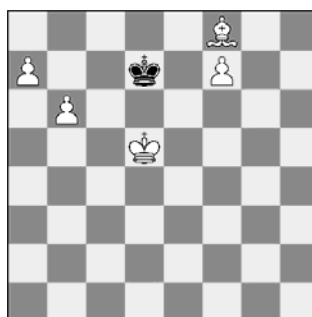
1. Bd6 Kc8 2. a8=Q+ Kd7 3. f8=N#

Rybka 3 vs. Fritz 8
2010

Fig. 2. The aesthetics of a human composition vs. chess engine match combination

(A) White to play and mate in 3 moves
1. Qg7 Ba7+ (d5/Bc7, Rh6#) **2. d4**
Bxd4+ (Bb8, d5#; d5/Bb6/Bc5, Rh6#)
3. Nxd4#
Aesthetic score: 3.95

(B) White to play and mate in 3 moves
1. Ke6 Nf7 (Na6/Nc6, Qxh8#; Kd8,
Qa7, Nf7, Rxb8#; Kf8, Rxb8#) **2.**
Qh8+ Nxh8 3. Rxb8#
Aesthetic score: 4.03

Fig. 3. Examples of the computer generated compositions using the experience approach

This does not mean, however, that the automatic composer can currently compete with the best human composers. Composition tournaments, and their judges, often have requirements and conventions that are not necessarily associated with beauty in the game but typically conflated with it. Such requirements or conventions (e.g. *no key move that restricts the enemy king's movement*) can be added as additional filters, if so desired, but these will likely significantly lower the number of successfully generated compositions. Like human composers, the automatic composer needs to be made 'aware' of these rules or it stands to have most of its compositions rejected.

The improvements obtained mainly in efficiency and somewhat in aesthetics using an experience table can be used to automatically generate *more* compositions that 'work' (or the same number in *less* time), and possibly more of reasonable quality.

This can be of value in terms of entertaining players and perhaps even aiding human composers (both amateurs and experts) by providing ideas they can further develop for their particular needs. Fig. 3 shows two comparable high-scoring examples of the automatically generated compositions from the pool of those generated by the HC and CG experience approaches; both the HC and CG are considered to be equivalent based on the experimental results shown in section 4. The main line is shown in bold whereas notable variations are shown in brackets. An experienced FIDE composition judge and solver, Michael McDowell, provided a detailed analysis of these two compositions – without being told they were composed by a computer or provided with the solutions – and was of the opinion that: "*In summary, to the experienced solver B has better content than A, but both score poorly for beauty and neither is of sufficiently high quality to be published in a reputable chess problem magazine.*"

The experience approach – suitably adapted – is also, in principle, applicable to other variants of the game like fairy chess and even Shogi. The only caveat is perhaps the requirement of a sizeable collection of human compositions of that type to gain 'experience' from; we would assume at least 1,000 combinations. The approach can be combined with other methods as described in section 2, possibly with improved results. Automatically composing *longer* problems (e.g. four or five-movers) – relatively rare in international chess – would likely be constrained by the strength and speed of the solving engine, and would require a suitably adapted aesthetics model.

6   Conclusion

Composing high quality chess problems requires considerable experience, knowledge, effort and time by human composers. Doing so computationally is therefore a challenge. Nevertheless, it is certainly possible to strictly *compose* problems – in principle, for any complex board game like chess – using a particular combination of techniques and technologies. Even so, many of them are computationally-intensive and therefore limit performance.

In this article, we have shown through experimentation that a simple 'experience table' can – at a very low computational cost – improve the automatic composition of chess problems in terms of efficiency and in some cases to a small degree, aesthetics as well. In order to do so, one first needs a sizeable database of compositions by humans in order to derive the piece-placement probability values; but beyond that, the compositions generated by the computer are sufficient as a source of 'experience'. A proper piece placement and position transformation strategy is also necessary to avoid convergence upon very similar-looking compositions. This idea, to the best of our knowledge, has never been tested before in the automatic composition of chess problems, but has now been shown to be a viable option. The output and entertainment potential of this technology can be considered reasonably significant.

Future work in this area would include, 1) looking at the effects of combining the various approaches of composing as described in section 2 and in this research to find the 'perfect mix' that provides the best compromise between efficiency and quality, 2) finding a way to apply the aesthetics model used at the 'knowledge level' so the automatic composer places pieces on the board with aesthetics in mind instead of just focusing on compositions that 'work', and 3) factoring in various composition

conventions, in such a way that does not significantly compromise composing efficiency or beauty, in order to be able to compete with the best human composers. Ultimately, the idea of a continuous feedback loop of 'experience' should be explored to see if a computer can learn and grow from its own composing experience.

References

1. Sezgin, F. (ed.): Book on Chess (Kitab al-shatranj): Selected Texts from al-Adli, Abu Bakr al-Suli and Others. Institute for the History of Arabic-Islamic Science, Johann Wolfgang Goethe University, Frankfurt am Main (1986)
2. Giddins, S.: Problems, Problems, Problems. ChessBase News (April 16, 2010), http://www.chessbase.com/newsdetail.asp?newsid=6261
3. Horvitz, E., Getoor, L., Guestrin, C., Hendler, J., Konstan, J., Subramanian, D., Wellman, M., Kautz, H.: AI Theory and Practice: A Discussion on Hard Challenges and Opportunities Ahead. AI Magazine 31(3), 103–144 (2010)
4. Hsu, F.-H.: Cracking Go. IEEE Spectrum 44(10), 50–55 (2007)
5. Ekbia, H.R.: Artificial Dreams: The Quest for Nonbiological Intelligence. Cambridge University Press, Cambridge (2008)
6. Hartmann, D.: Human Superiority Restored. ICGA Journal 33(3), 150 (2010)
7. Boden, M.A.: Computer Models of Creativity. AI Magazine 30(3), 23–34 (2009)
8. Manaris, B., Roos, P., Penousal, M., Krehbiel, D., Pellicoro, L., Romero, J.: A Corpus-Based Hybrid Approach to Music Analysis and Composition. In: The 22nd Conference on Artificial Intelligence (AAAI 2007), pp. 839–845. AAAI Press, Vancouver (2007)
9. Rigau, J., Feixas, M., Sbert, M.: Conceptualizing Birkhoff's Aesthetic Measure Using Shannon Entropy and Kolmogorov Complexity. In: The Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging, pp. 105–112. Eurographics Assoc., Banff (2007)
10. Iqbal, M.A.M.: A Discrete Computational Aesthetics Model for a Zero-sum Perfect Information Game. Ph.D. thesis. Faculty of Computer Science and Information Technology. University of Malaya, Kuala Lumpur, Malaysia (2008), http://metalab.uniten.edu.my/~azlan/Research/pdfs/phd_thesis_azlan.pdf
11. Iqbal, A.: Aesthetics in Mate-in-3 Combinations, Part I: Combinatorics and Weights. ICGA Journal 33(3), 140–148 (2010)
12. Schlosser, M.: Computers and Chess Problem Composition. ICCA Journal 11(4), 151–155 (1988)
13. Schlosser, M.: Can a Computer Compose Chess Problems? In: Beal, D.F. (ed.) Advances in Computer Chess, vol. 6, pp. 117–131. Ellis Horwood Ltd., Chichester (1991)
14. Haworth, G.M.: Chess Endgame Knowledge Advances. ICGA Journal 33(3), 149 (2010)
15. Noshita, K.N.: A Note on Algorithmic Generation of Tsume-Shogi Problems. In: The Game Programming Workshop 1996, Kanagawa, Japan, pp. 27–33 (1996)
16. Hirose, M., Matsubara, H., Itoh, T.: The Composition of Tsume-Shogi Problems. In: Advances in Computer Chess, vol. 8, pp. 299–318. Universiteit Maastricht, the Netherlands (1996)
17. Watanabe, H., Iida, H., Uiterwijk, J.W.H.M.: Automatic Composition of Shogi Mating Problems. In: Games in AI Research, pp. 109–123. Universiteit Maastricht, the Netherlands (2000)

18. Horiyama, T., Ito, H., Iwama, K., Kawahara, J.: Enumeration of Tsume-Shogi Diagrams by the Reverse Method. In: The International Conference on Informatics Education and Research for Knowledge-Circulating Society, Kyoto, Japan, pp. 193–196 (2008)
19. HaCohen-Kerner, Y., Cohen, N., Shasha, E.: An Improver of Chess Problems. Cybernetics and Systems 30(5), 441–465 (1999)
20. Fainshtein, F., HaCohen-Kerner, Y.: A Chess Composer of Two-Move Mate Problems. ICGA Journal 29(1), 32–39 (2006)
21. Fainshtein, F., HaCohen-Kerner, Y.: A Deep Improver of Two-move Chess Mate Problems. Cybernetics and Systems 37(5), 443–462 (2006)
22. Bilalić, M., McLeod, P., Gobet, F.: Inflexibility of Experts – Reality or Myth? Quantifying the Einstellung Effect in Chess Masters. Cognitive Psychology 56(2), 73–102 (2008)
23. Walls, B.P.: Beautiful Mates: Applying Principles of Beauty to Computer Chess Heuristics. M.Sc. dissertation. University of Sussex, UK (1997)
24. Albrecht, H.: How Should the Role of a (Chess) Tourney Judge Be Interpreted? The Problemist 7, 21–218 (2000)
25. Iqbal, A.: Aesthetics in Mate-in-3 Combinations, Part II: Normality. ICGA Journal 33(4), 202–211 (2010)
26. Thomson, K.: Retrograde Analysis of Certain Endgames. ICCA Journal 9(3), 131–139 (1986)
27. Smullyan, R.: Chess Mysteries of Sherlock Holmes: Fifty Tantalizing Problems of Chess Detection. Random House Puzzles, New York (1994)
28. Lipton, M., Matthews, R.C.O., Rice, J.: Chess Problems: Introduction to an Art. Citadel Press, New York (1965)
29. Iqbal, A.: Evaluation of Economy in a Zero-sum Perfect Information Game. The Computer Journal 51(4), 408–418 (2008)
30. Sadikov, A., Bratko, I.: Search Versus Knowledge Revisited Again. In: van den Herik, H.J., Ciancarini, P., Donkers, H.H.L.M(J.) (eds.) CG 2006. LNCS, vol. 4630, pp. 172–180. Springer, Heidelberg (2007)

Appendix