

Origin of chess - a phylogenetic perspective /

Alex R. Kraaijeveld

Board games are like plant and animal species in that they can evolve and give rise to new forms. This leads to an important similarity between board games and biological species (Eagle 1995, 1998; Kraaijeveld 1999): in both cases a group of similar-looking 'species' most likely owe their similarity to their being descended from a single common ancestor. Obviously, board games can also go extinct and they even have their 'fossils': games that are not played anymore today, but are known (completely or partly) from historical sources. The Egyptian games of Tau and Senet, the Chinese Liubo and the larger Shogi (Japanese chess) variants are just a few examples of such fossils and Tai-Kyoku Shogi, with a board of 36x36 squares and more than 400 pieces per side (Umebayashi 1997), can be considered the board games equivalent of a brontosaurus fossil.

A field of biology, called phylogeny, uses a range of analytical methods in reconstructing the evolution of groups of plants or animals. These phylogenetic techniques are widely accepted tools in evolutionary biology (with a vast body of literature; see Kitching et al (1998) for an excellent introduction to the subject) and have also been successfully used in research into the evolution of languages (Ruhlen 1987). Phylogenetic analyses are based on two important concepts. First of all, characters of a species can occur in one of two states: primitive or derived (the technical terms are plesiomorphic and apomorphic, respectively). The primitive state is the older of the two, evolving into the derived state at some point in time. Shared derived characters point to relatedness, whereas shared primitive characters do not. A biological example of a shared derived character is the possession of feathers as body covering. Feathers is one of the traits that makes an animal a bird and all birds are generally considered to form a closely related group of animals. Feathers evolved from scales, and therefore, having scales as body covering is the primitive state. Both a herring and a lizard have scales, but sharing this primitive character does not make them closely related. The second important concept is that of parsimony. The parsimony principle states that it is unlikely (but not impossible!) that the same change from a primitive to a derived character state occurs in two species independently of each other. Therefore, in reconstructing the evolution of a group of species, the analysis searches for the evolutionary tree with the lowest number of such independent similar changes.

Phylogeny and the ancestor of Chess

Probably the most hotly debated issue among chess historians is where and when chess originated. Intricately linked to this is the question what the ancestral form of chess, which has given rise to a multitude of variants around the world, looked like.

The traditional view, advocated by Hyde (1694), Forbes (1860), Murray (1913), Bhatta (1994) and many others is that chess originated in India, possibly in the late 6th or early 7th century. Although different views have been around for a long time (see Li (1998) for a detailed overview), the last few decades have seen opposition against the traditional "out-of-India" view become much stronger: Needham (1962), Bidev (1986),

Li (1998) and several others see the origin of chess in China, either earlier in the 6th century or around 200 BC. Debate at the moment is largely, but not exclusively, polarised between the “out-of-India” and the “out-of-China” theories (but see Josten (1995) for a challenge to the whole idea of a single place of origin).

Basically, three sources of information could shed light on where and when chess originated: literary sources, archaeological finds and the information that is contained within the game itself (i.e. shape of the game, pieces and their moves, rules of play, etc). Evidence from early literary sources is open to several interpretations and has, so far, not been able to firmly pinpoint the birthplace of chess. The oldest archaeological finds that are generally considered to be chess pieces are the Afrasiab pieces from Persia (Calvo 1998). As they are dated to the 8th (or possibly 7th) century, they are probably too young to be regarded as pieces of the ancestral form of chess. Interestingly, the earliest solid pieces of both literary and archaeological evidence for the existence of a chess-like game all seem to be neither Indian nor Chinese, but Persian (Calvo 1994).

The third source of information, the game itself, is by far the richest at the moment. Many chess historians over the centuries have isolated similarities between variants to support their specific claim as to the identity of the ancestor of chess. A problem is that, in general, humans have a natural tendency to focus on pieces of information that support a certain hypothesis while paying less attention to those that contradict it. In other words, a step forward in using the available information contained within the game is to combine all the information in an objective way, without any preconceived ideas, and phylogenetic techniques offer exactly this possibility.

A chess analogue to the scales - feathers example mentioned earlier is the move of the queen-type piece: its move in medieval chess is primitive, its modern move derived. It is important to realise that the absence of something is not necessarily the primitive state with its presence consequently suggesting relatedness. If two chess variants both have a knight as one of their pieces, that does not mean they are closely related. Virtually all chess variants have a knight among their pieces, so the knight was almost surely present in the ancestral form of chess; therefore the presence of a knight is almost surely the primitive state, its absence the derived state.

Because shared derived characters suggest relatedness whereas shared primitive characters do not, it is crucial that for each character it is known which state is derived and which state is primitive. If the direction of character change is entered in the analysis in the wrong way, relatedness will be inferred from shared primitive characters. In other words, using the wrong ancestor (whose characters are by definition all in a primitive state), will result in the wrong evolutionary tree (“garbage in - garbage out”). However, this can be used as a tool in looking at the origin of chess from an evolutionary viewpoint. Phylogenetic analyses can be done with a number of hypothetical ancestors and if an evolutionary tree resulting from such an analysis shows considerable discrepancies with what is known from historical sources, it is very likely that the ancestor used in that particular tree is quite different from the actual ancestor of chess.

Data & Methodology

A selection of 40 chess variants was made, representing the full range in this group of board games both in terms of geography and time. These 40 variants belong to a number of largely geographical subgroups which contain variants that are generally considered closely related (the main reason for this subgrouping is to make it easier to determine how well the tree resulting from a particular analysis matches historical knowledge): Indian (five variants, including the 2-sided Chaturanga and the 4-sided dice-form Chaturaji), South-East Asian (five variants), Central Asian (three variants), Arabo-Persian (eight variants, from the basic Shatranj to more elaborate variants like Timur's chess and Grande Acedrex; this group also includes Senterej and Samantsy from Ethiopia and Madagascar, respectively), early medieval and later European (seven variants in total, from early medieval forms to the modern FIDE-version, here referred to as Western chess) and the New World (five variants, of which Aleut chess could arguably also be grouped with the Central Asian variants (Murray 1913; Jochelson 1933), but the reliability of the information on this variant is doubted (Kraaijeveld 1997)). Because of the controversy surrounding the origin of Japanese chess (i.e. descending from a Chinese or from a South-East Asian branch; Masukawa 1994) Japanese and Chino-Korean variants are treated as two small groups (with two and three variants, respectively), separate from the other Asian variants. The only Japanese variants included here are modern Shogi and the earliest form of Shogi: the 8x8 version of Heian Shogi (Masukawa 1994). The reasons for not including any of the large Shogi variants are first of all that due to the large numbers of exotic pieces, they are so different from all other chess variants that they can not shed any light on the deeper-lying question of the ancestor of chess, and secondly, their evolutionary relationship has been investigated in an earlier publication (Kraaijeveld 1999). The early form of Xiangqi (Chinese chess) in the Chino-Korean group is the one described by Himly (1871). Finally, two fictional variants are included: Jetan (from Edgar Rice Burroughs' novel *The Chessmen of Mars*) and Federation chess (the 3-dimensional variant featuring in the sci-fi series and movies *Star Trek*); the reasons for including both these fictional variants will be explained below.

For each of the 40 variants, the state of 109 characters was scored as '0' or '1' (i.e. no/yes or absent/present) or '?' if unknown or irrelevant, resulting in a data-matrix of 4360 '0's, '1's and '?'s. These characters are both 'morphological' (i.e. physical characteristics of the board and pieces, presence/absence of pieces, the initial set-up) and 'physiological' (moves of the various pieces, promotion rules, rules relating to the end of the game); the full list of characters used is given in the appendix. Data on most chess variants were taken from Murray (1913) and Pritchard (1994), with additional information from Jochelson (1933; on Aleut chess) and Ivanova (1994; on Tuvinian chess).

The analyses were performed with the PHYLIP-package, which can be obtained from the Internet (<http://evolution.genetics.washington.edu/phylip.html>). Because calculating phylogenetic trees is sensitive to the order in which species are entered, each analysis was done 1000 times, with the order of variants randomised. Each set of 1000 runs resulted in a (varying) number of equally parsimonious trees, which were subsequently merged into one consensus tree per analysis.

Several variants have been suggested as the ancestor of all other chess variants. Traditionally, the ancient Indian Chaturanga is seen as at least very close to the ancestor of chess, but the 4-sided Indian dice-variant Chaturaji has also been regarded as ancestral at various points in time. More recently, Li (1998) presented a hypothetical reconstruction of the original form of Xiangqi (referred to as 'proto-Xiangqi' here). Other ideas on the identity of the ancestor have been suggested, but as no full details on the pieces, moves and other rules of these hypothetical ancestors were given, it is not possible to include them here. Therefore, phylogenetic analyses are performed using Chaturanga, Chaturaji and proto-Xiangqi as the hypothetical ancestor. To illustrate that using the wrong ancestor gives demonstrably wrong results, an analysis was first done with as ancestor a variant of which we can be absolutely sure that it is not the correct one: Federation chess (the 3-dimensional fictional variant from Star Trek). In each of the analyses, the choice of the ancestor by definition makes the character states of that variant primitive for that particular analysis.

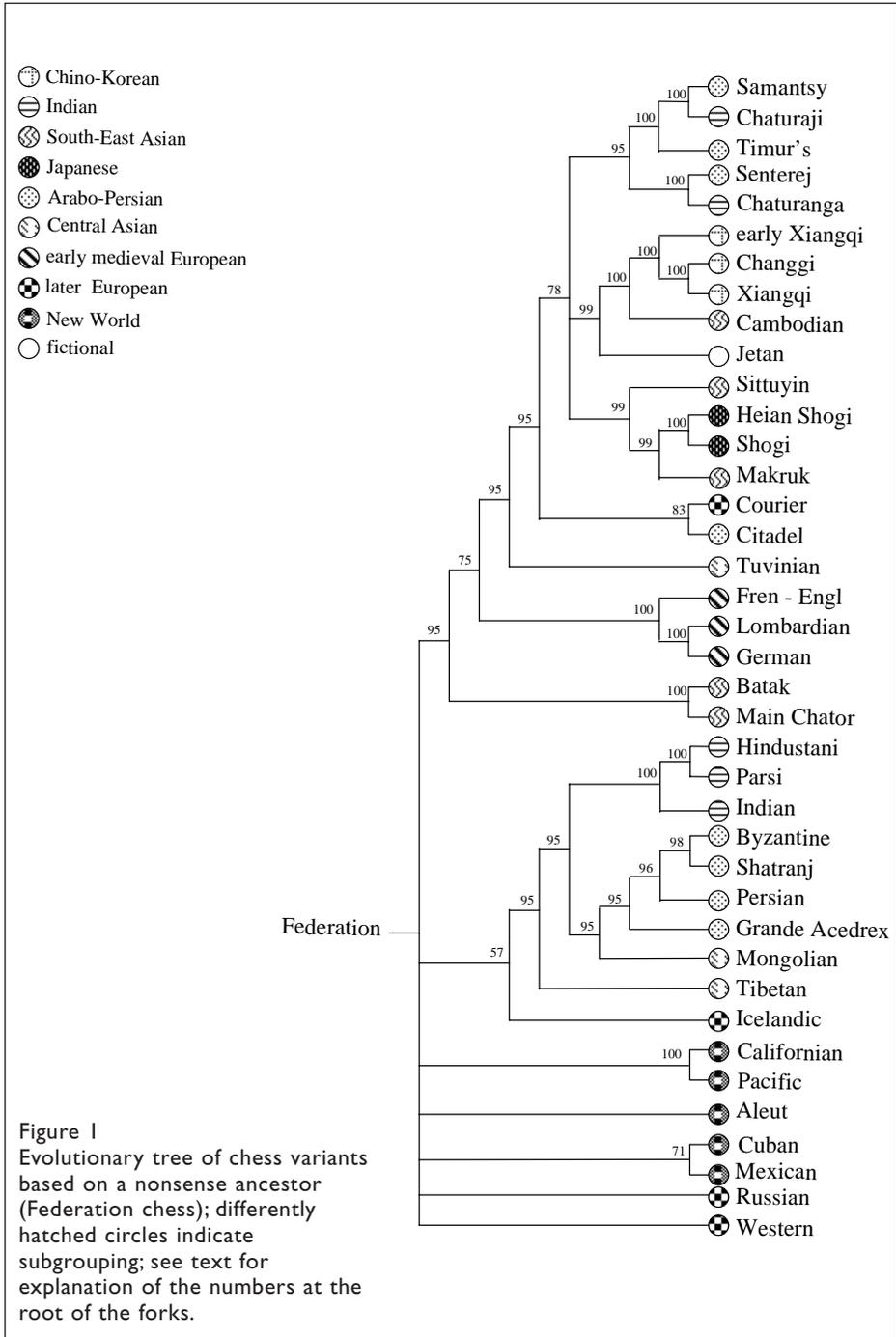
Results

The evolutionary trees resulting from the four analyses (one with a nonsense ancestor and three with a hypothetical ancestor) are given in figures 1-4. In all figures, the numbers at the root of each fork indicate the percentage of equally parsimonious trees that contained that fork; in other words, values close to 100 indicate that the analysis gives strong support for the variants in that fork being related; forks with less than 50% occurrence were considered unresolved and collapsed to the next higher level. To help visually assess the likelihood of the hypothetical ancestor being close to the actual one, each variant name is preceded by a circle with a hatching pattern that indicates which of the subgroups mentioned earlier that particular variant belongs to.

Figure 1 illustrates the effects of using a wrong ancestor. Although some groups indeed come out properly (like the Chino-Korean variants), there are quite a number of discrepancies with historical knowledge. Chaturanga and Chaturaji appear to be only distantly related to the other Indian variants. The Arabo-Persian variants end up scattered through the tree rather than in one group. The three early medieval European variants (Lombardian, German and French-English chess) do come out as one group, but unrelated to the later European variants, which in reality descended from them. In fact, several European variants (plus the New World variants) come out isolated and unrelated to anything else.

The evolutionary tree with Chaturanga as ancestor is given in figure 2. Almost all groups come out intact, with the Arabo-Persian variants basic to the early medieval European variants, which in turn are basic to the later European and New World variants; this is precisely in accordance with historical knowledge. The only 'strange' group is the one combining Timur's chess and the two Arabo-Persian variants from African soil (Senterej and Samantsy) with Chaturaji and Jetan.

Using Chaturaji as ancestor (figure 3) leads to several groups coming out more scattered than with Chaturanga as ancestor: all main branches now contain at least one Arabo-Persian variant and both the early medieval European and the Central Asian



groups are split up, with Tuvinian chess coming out quite distant from Mongolian and Tibetan chess.

The final tree (figure 4) uses proto-Xiangqi as ancestor. Again the Arabo-Persian variants are scattered around the tree; Byzantine chess, a circular form of Shatranj, comes out as only distantly related to its direct ancestor. Like in the Chaturaji-based tree the early medieval and Central Asian variants are split up. Proto-Xiangqi as ancestor also leads to the 'strange' grouping of Timur's chess and the two African Arabo-Persian variants with Chaturaji and Jetan.

Discussion

From the two trees with an Indian ancestor, the one with Chaturanga gives a better match with historical knowledge than the one with Chaturaji. This suggests that the 4-sided dice-game Chaturaji is less likely to be the ancestor of chess than Chaturanga. This is in accordance with historical sources, which suggest that the 4-sided dice-game was an experiment of a much later date than the 2-sided game. The lower probability of Chaturaji as the ancestor of chess can not solely be caused by it being a dice-controlled game. Whether play is controlled by dice or not is only one of 109 characters; the analyses take the full set of characters into account, without focussing on a single one.

Comparison of the Chaturanga-based tree (being the better of the two "out of India" trees) with the proto-Xiangqi-based tree shows that the former is in better agreement with historical events than the latter. In other words, the phylogenetic analyses performed here suggest that the ancestor of the range of chess variants used here is more similar to Chaturanga than to Xiangqi. This does not necessarily mean that chess originated in India. The Chaturanga-like ancestor suggested by the analyses could itself have come from somewhere else or be a descendant of an older form that originated outside India. Schädler (1999) presents details of a form of the Royal Game of Ur from the 2nd century BC. The game has five different pieces (birds; the moves of which are determined by dice), which makes it the only game besides and prior to chess variants with a similar degree of piece differentiation. Including this game as a potential ancestor of chess in a phylogenetic analysis is not possible with the present state of knowledge, however. Phylogenetic analyses can look no further back in time than the most recent common ancestor of a group of 'species'. For an analysis which could include this Ur game as an ancestor a number of games that are descendants of the Ur game and not chess variants would be needed. Another point is that phylogenetic analyses in general assume that all species included descend from a single common ancestor. Although this is generally regarded to hold for chess, recent ideas doubt this and see different forms of proto-chess come into being at several places simultaneously along the Silk Road by amalgamations of elements from various other games (Josten 1995).

An important difference between biological species and board games is that 'borrowing' of elements between board games is much easier than transfer of genetic material between biological species. This is a potential problem in the application of phylogenetic techniques to board games research, because it weakens the parsimony principle. A prime example of borrowing between chess variants is the addition of the Western rook

and bishop to Shogi. Usually, however, analyses will be based on many characters (over 100 in the present case) which makes the effect of a low number of borrowings on the overall outcome negligible (Shogi never comes out as closely related to Western chess in the analyses, despite the presence of this borrowing). Also, application of phylogenetic techniques in linguistic research has been very successful, even though borrowing can also happen quite easily between languages.

Timur's chess and Jetan illustrate a limit to the use of phylogenetic techniques in this field. Although Timur's chess is known from historical sources to be derived from Shatranj, it is quite different from its ancestor and has few shared derived traits with Shatranj. This is most likely the reason for it coming out quite distant from the other Arabo-Persian variants in all analyses. In the case of the fictional Jetan, this is even more extreme: Jetan was obviously made up by a man who had a knowledge of Western chess (and thus, in a way, Western chess is its direct ancestor), but it is so different in most of its characters from any other variant that it never ends up close to Western chess in the trees, but in various strange places, depending on the ancestor chosen. On a more general level, this indicates that phylogenetic techniques will have difficulty dealing with board game variants that come into being by a sudden big leap of imagination. Such sudden big chances from an ancestor to its direct descendent is something that can not happen in biological species or even languages.

The use of phylogenetic techniques in board games research is not meant to replace historical (literary or archaeological) research, but as a complementary new method and a way to look at the evolution of board games from a very different perspective. A phylogenetic analysis on a group of 15 Shogi variants (Kraaijeveld 1999) suggested a multiple origin of the larger variants rather than a single process of variants getting bigger and bigger over time. Historical sources are necessary to provide evidence for this theory, for instance by showing that some of the larger variants originated independently in different Buddhist monasteries; Buddhist monks are often considered the inventors of at least some of the larger Shogi variants (Hodges 1982a, 1982b, 1983).

In this article, looking at the question of the origin of chess from a phylogenetic perspective lends support to the theory that the ancestor of the many chess variant known through time and from around the world was similar to Chaturanga. The one unexpected group in the Chaturanga-based tree (the one grouping Samantsy and Senterej with Chaturaji), however, suggests that some of the characters of this ancestor were not completely identical to Chaturanga. On a more general level, to evaluate the usefulness of phylogenetic techniques to board game research it will be necessary to apply them to other groups of board games; the large group of mancala games seems a promising candidate for this.

Acknowledgements

My thanks to Mark Fellowes and Ulrich Schädler for critical comments on previous versions of the manuscript.

References

- Bhatta, C.P. 1994. *Origin and Genesis of Chess*. Seevetal.
- Bidev, P. 1986. *Stammt Schach aus Alt Indien oder China?* Igalo Press.
- Calvo, R. 1994. Some flashes of lightning into the “impenetrable darkness” of chess origins, In: *The Chess Collector*, IV: 14-28.
- Calvo, R. 1998. Some facts to think about. In: Josten, G. & Meissenburg, E. (eds.). *Where did chess originate? Dissenting votes*, Seevetal: 11-13.
- Eagle, V.A. 1995. On some newly described Mancala games from Yunnan province, China, and the definition of a genus in the family of Mancala games. In: A.J. de Voogt (ed.), *New approaches to board games research: Asian origins and future perspectives*, IIAS, Leiden: 48-62.
- Eagle, V.A. 1998. On a phylogenetic classification of Mancala games, with some newly recorded games from the “southern Silk Road”, Yunnan province, China, In: *Board Games Studies*, 1: 50-68.
- Forbes, D. 1860. *The History of Chess*. Allen, London.
- Himly, K. 1871. The Chinese game of Chess as compared with that practised by Western nations, In: *Journal of the North-China Branch of the Royal Asiatic Society for 1869 and 1870*: 107-121.
- Hodges, G.F. 1982a. *Dai-Dai Shogi and how to play it*. The Shogi Association, Bromley.
- Hodges, G.F. 1982b. *Maka-Dai-Dai Shogi and how to play it*. The Shogi Association, Bromley.
- Hodges, G.F. 1983. *Tenjiku Shogi*. The Shogi Association, Bromley.
- Hyde, T. 1694. *De Ludis Orientalibus libri duo*. Oxford.
- Ivanova, N. 1994. Chessmen from Siberia, In: *CCI St Petersburg Conference Proceedings*.
- Jochelson, W. 1933. *History, Ethnography and Anthropology of the Aleut*. Carnegie Institution, Washington.
- Josten, G. 1995. *Der Läufer war eine Dame*, Maintal, Rochade Europa.
- Kitching, I.J., Forey, P.L., Humphries, C.J. & Williams D.M. 1998. *Cladistics. The theory and practice of parsimony analysis*. Oxford.
- Kraaijeveld, A.R. 1997. Yakutat and Aleut chess: indigenous Alaskan chess variants?, In: *Variant Chess*, 3: 113-114.
- Kraaijeveld, A.R. 1999. Phylogenetic analysis of Shogi variants, In: *Variant Chess*, 4: 56-58.
- Li, D.L. 1998. *The Genealogy of Chess*. Premier Publishing Company, Bethesda.
- Masukawa, K. 1994. The origin of Japanese chess, In: *Variant Chess*, 2: 100-104.
- Murray, H.J.R. 1913. *A History of Chess*. Oxford.
- Needham, J. 1962. The Magnet, Divination, and Chess. In: *Science and Civilisation in China*. Vol. 4, Pt. 1, Cambridge: 314-334.
- Pritchard, D.B. 1994. *The Encyclopedia of Chess Variants*. Games & Puzzles Publications, Godalming.
- Ruhlen, M. 1987. *A Guide to the World's Languages*. Vol. 1: *Classification*. Stanford.
- Schädler, U. 1999. Vom 20-Felder-Spiel zum Würfelierschach?, In: *Board Games Studies*, 2: 144-148.
- Umabayashi, I. 1997. *Sekai no Shogi*. Shogi Tengokusha.

Appendix

The full list of characters used in the analyses

- 2 players
 board 2-dimensional
 wide board
 deep board
 circular board
 play on squares
 squares checkered
 palace
 river
 citadels
 ashtapada markings
 long diagonals
 movable parts
 dice involved
 shape of major pieces
 shape of pawns
 different name for opposing pieces
 pawn row advanced
 pawn row continuous
 other pieces in pawn row
 one row of pawns
 one row of major pieces
 rook-type piece; in corner
 rook-type piece; limit to 2 squares
 rook-type piece; only forward
 knight; present
 knight; among main pieces
 knight; move all directions
 knight; can leap
 bishop-type piece; can leap
 bishop-type piece; extended knight move
 bishop-type piece; silver general move
 bishop-type piece; king move
 bishop-type piece; modern bishop move
 queen-type piece; present
 queen-type piece; doubled
 queen-type piece; additional rook move
 queen-type piece; modern queen move
 queen-type piece; additional knight move
 queen-type piece; golden general move
 queen-type piece; king move
 queen-type piece; limit to 2-3 squares
 queen-type piece; gryphon move
 queen-type piece; can capture
 queen-type piece; capture all directions
 queen-type piece; joyleap
 queen-type piece; escape
 king; in main row
 king; facing opposing king
 king; both orthogonal and diagonal move
 king; 3 steps
 king; knight move
 king; leap on first move
 king; direct facing allowed
 king; simultaneous move with queen-type piece
 castling
 pawns; capture diagonal
 pawns; capture en passant
 pawns; sideways move
 pawns; sideways after crossing river
 pawns; initial double step
 pawns; all initial double step
 bat present
 camel present
 castle present
 cavalier present
 courtier present
 fortress present
 dabbaba present
 courier present
 wazir present
 nobleman present
 counsellor present
 giraffe present
 talia present
 lion present
 unicorn present
 canon present
 additional rook & bishop present
 dwar present
 main pieces change direction while moving
 knight & elephant transposition
 full marshalling
 opening moves alternating
 ritual first moves
 captured pieces disappear from game
 main pieces can promote
 pawns can promote
 promotion at back row
 promotion at diagonal
 promotion instantaneous
 promotion after joyleap
 promotion optional
 promotion to file piece
 promotion to lost piece
 promotion to piece of choice
 indication of promoted pieces
 capturing king ends game
 capturing princess ends game
 bare king loses
 bare king is draw
 bare king changes moves
 stalemate is win
 stalemate is draw
 stalemate is illegal
 knight can mate
 sequential mating
 pawns must be gone for win
 victim replaces king when mated