Deeper model endgame analysis

Deeper model endgame analysis

R.B. Andrist and G.M^{c}C. Haworth

rba_schach@gmx.ch; g.haworth@reading.ac.uk

Abstract

A reference model of Fallible Endgame Play has been implemented and exercised with the chess-engine WILHELM. Past experiments have demonstrated the value of the model and the robustness of decisions based on it: experiments agree well with a Markov Model theory. Here, the reference model is exercised on the well-known endgame KBBKN.

Key words: chess, endgame, experiment, fallibility, Markov, model, theory

1 Introduction

In previous papers [6,7], a reference model of fallible endgame play has been defined in terms of a spectrum of Reference Endgame Players (REPs) R_c. The REPs are defined as choosing their moves stochastically from an endgame table (EGT), using only the values and depths of successor positions.

Here, we survey and compare existing experimental and theoretical results, and report on the latest findings with the familiar, complex endgame KBBKN. In Section 2, we revisit the basic concepts and theory of the REP model, while in Section 3, we describe the REP implementation in WILHELM [1]. In Sections 4-7, we review past experiments, compare experiment and theory, and introduce the KBBKN results. Section 8 summarises and notes some questions arising from this work.

2 The reference endgame player model

A nominated endgame, e.g., chess' KQKR, is considered to be a system with a finite set of states $\{s_i\}$ numbered from 0 to ns-1.[1] Each state $s(val, d)$ is an *equivalence class* of positions of the same theoretical value val and depth d. Higher-numbered states are assumed to be less attractive to the side to move, which is taken to be White. Thus, for KQKR with the DTC[2] metric, we have maxDTCs (1-0) n_W =31, (0-1) n_B = 3, and ns = 37 states in total:

- s_i, $i = 0$: a 1-0 win, i.e. for White, not requiring a winner's move[3],
- s_i, $1 \le i \le 31$: 1-0 wins of depth i,
- s_i, $i = 32$: theoretical draw, either in the endgame or a subgame,
- s_i, $33 \le i \le 35$: 0-1 wins, i.e. for Black, of depth 36-i
- s_i, $i = 36$: a 0-1 win not requiring a winner's move.

The REP R_c in position P chooses stochastically from moves which each have a probability proportional to a Preference[4], $S_c(s_s[val_s, d_s])$, where s is the move's destination state with

[1] For convenience, Appendix A summarises the key acronyms, notation and terms.
[2] DTC ≡ DTC(onversion) ≡ Depth to Conversion, i.e. to mate and/or change of material.
[3] i.e., mate, achieved conversion to won subgame, or loser forced to convert on next move.

theoretical value val_s and win/loss depth d_s. Each move-choice by R_c is independent of previous move-choices. We require that $\{R_c\}$ is a spectrum of players, ranging linearly from the metric-infallible player R_∞ via the random player R_0 to $R_{-\infty}$, the anti-infallible player. To ensure this, the function $S_c(s[val, d])$ is required to meet some natural criteria, as described more fully and formally in [7] and in Appendix B.

Here, we choose, as an $S_c(s_s[val, d])$ function meeting those criteria:

$S_c(s_s[win, d]) \equiv (d + \kappa)^{-c}$ with $\kappa > 0$ to ensure that S_c is finite,
$S_c(draw) \equiv S_c(win, n_1) \equiv S_c(loss, n_2)$ with $n_1 > n_W$ and $n_2 > n_B$
$S_c(s_s[loss, d]) \equiv \lambda \cdot (d + \kappa)^c$, λ being defined by n_1 and n_2 above.

This ensures, as required, that R_0 prefers no move to any other, that R_c with $c > 0$ prefers better moves to worse moves, and that as $c \to \infty$, the R_c increase in competence and tend to infallibility in terms of the chosen metric.

Although the R_c have no game-specific knowledge, the general REP model allows moves to be given a prior, ancillary, weighting v_m based on such considerations [9]. Thus, $v_m = 0$, as used in this paper, prevents a move being chosen and $v_m > 1$ makes it more likely to be chosen.

The probability $T_c(i)$ of moving from a position to state s_i is therefore:

$$T_c(i) \equiv S_c(s_i) \cdot \sum_{\text{moves\_to\_state\_}i} v_m \,/\, \sum_{\text{all\_moves}} v_m \cdot S_c(s_{move})$$

3 Implementing the REP model

The first author has implemented in WILHELM [1] a subset of the REP model which is sufficient to provide the results of this paper. Ancillary weightings v_m are restricted to 1 and 0. $v_m = 0$ is, if relevant, applied to all moves to a state s rather than to specific moves: it can be used to exclude moves losing theoretical value, and/or to emulate a search horizon of H moves, within which a player will win or not lose if possible. WILHELM offers five agents based on the REP model: these are, as defined below, the *Player*, *Analyser*, *Predator*, *Emulator* and *Predictor*. A predefined number of games may be played between any two of WILHELM, *Player*, *Predator*, *Emulator* and an infallible player with endgame data. WILHELM also supports the creation of Markov matrices, see Section 5.

3.1 The Player

The *Player* is an REP R_c of competence c, and therefore chooses its moves stochastically using a validated (pseudo-)random number generator in conjunction with the function $S_c(val, d)$ defined earlier.

[4] For convenience and clarity, the Preference Function $S_c(val_s, d_s)$ may be signified by the more compact notations $S_c(val, d)$ or merely $S_c(s)$ if the context allows.

3.2 The Analyser

Let us imagine that an unknown fallible opponent is actually going to play as an R_c with probability $p(x)\cdot\delta x$ that $c \in (x, x + \delta x)$: $\int p(x)\mathrm{d}x = 1$. The *Analyser* attempts to identify the actual, underlying c of the R_c which it observes. For computational reasons, the Analyser must assume that c is a value from a finite set $\{c_j\}$ and that $c = c_j$ with initial probability $pc_{0,j}$. Here, the c_j are regularly spaced in $[c_{min}, c_{max}]$ as follows:

$c_{min} = c_1$, $c_j = c_1 + (j\text{-}1)\cdot c\delta$ and $c_{max} = c_1 + (n\text{-}1)\cdot c\delta$, i.e. $c = c_{min}(c\delta)c_{max}$.

The notation $c = c_{min}(c\delta)c_{max}$ is used to denote this set of possible values c. The initial probabilities $pc_{0,j}$ may be $1/n$, the usual 'know nothing' uniform distribution, or may be based on previous experience or hypothesis. They are modified, given a move to state s_{next}, by Bayesian inference [4]:

$T_j(next)$ = Prob[move to state $s_{next} \mid c = c_j$], and
$pc_{i+1,j} = pc_{i,j}\cdot T_j(next) / \sum_k [pc_{i,k}\cdot T_k(next)]$.

Thus, the new Expected[c] = $\sum_j pc_{i+1,j}\cdot c_j$.

In Subsection 4.1 below, we investigate what values should be chosen for the parameters c_{min}, $c\delta$ and c_{max} so that the errors of discrete approximation are acceptably small.

3.3 The Predator

On the basis of what the *Predator* has learned from the *Analyser* about its opponent, it chooses its move to best challenge the opponent, i.e., to optimise the expected value and depth of the position after a sequence of moves. As winning attacker, it seeks to minimise expected depth; as losing defender, it seeks to maximise expected depth. In a draw situation, it seeks to finesse a win. Different moves by the predator create different sets of move-choices for the fallible opponent. These in turn lead to different expectations of theoretical value and depth after the opponent's moves. The predator implementation in WILHELM chooses its move on the basis of only a 2-ply search. It may be that deeper searches will be worthwhile, particularly in the draw situation.

3.4 The Emulator

The *Emulator* E_c is conceived as a practice opponent with a 'designer' level of competence tailorable to the requirements of the practising player. An REP R_c will exhibit an apparent competence c' varying, perhaps widely, above and below c because it chooses its moves stochastically. In contrast, the *Emulator* E_c chooses a move which exhibits to an Analyser an apparent competence c'' as close to c as possible.

The reference Analyser is defined as initially assuming the Emulator is an R_x, $x = 0(1)2c$, where $x = x_j$ with initial probability $1/(2c+1)$. The *Emulator* E_c therefore opposes a practising player with a more consistent competence c than would R_c, albeit with some loss of variety in its choice of moves. The value c can be chosen to provide a suitable challenge in the practice session. The practising player may also have their apparent competence assessed by the *Analyser*.

3.5 The Predictor

The *Predictor* is advised of the apparent competence c of the opponent. It then predicts how long it will take to win, or what its chances are of turning a draw into a win, using data from an Analyser and from a Markov Model [4] of the endgame. This model is defined in Section 5 below.

4 A review of previous experiments

The first use of the REP model and WILHELM [6,7] was to study the two famous Browne-BELLE KQKR exhibition games [5,10]. Browne's apparent competence c was assessed by an Analyser, and BELLE's moves as Black were compared with the decisions of a Predator using the Analyser's output.

Browne's apparent c was approximately 19, the highest figure so far measured in a fallible player. In comparison, Bronstein [11] and Timman [3] have both measured in at around c = 15 when attacking in KBBKN endgames.
Six choices had to be made to effect the numerical analysis:

- $c_{min} = 0$, $c\delta = 1$, $c_{max} = 50$; $\kappa = 0^+$ (i.e., arbitrarily small, effectively zero)
- all c_j were deemed equally likely,
- metric = DTC.

It was natural to begin by testing the effect of these six choices in the next experiments [8].The aim was to examine the robustness of the Analyser's perception of Browne's apparent capability c, and any effect on the Predator's choice of moves.

4.1 The effect of numerical analysis choices

To test the choice of $c\delta$, Browne-BELLE game 1 was reanalyzed with:

$c_{min} = 0$, $c_{max} = 50$, $\kappa = 1$, and $c\delta$ in turn set to 0.01, 0.1, 1, 2, 5 and 10.

It may be shown the Analyser's Bayesian calculation is a discrete approximation to the integral of a Riemann-integrable function. Therefore, the theory of integration guarantees that this calculation will converge as $c\delta \to 0$. We judge that the error is ignorable with $c\delta$ = 1 and that no smaller $c\delta$ is needed. Similarly, the calculation converges as $c_{min} \to -\infty$ and $c_{max} \to \infty$. Given that Browne appeared to have a c of approximately 20, the choices of c_{min} = 0 and c_{max} = 50 had an insignificant effect on accuracy. It seems reasonable to assume that the opponent will demonstrate positive skill, and that a $c_{max} \approx 2.5 \times$ actual c should be appropriate. Of course, while the opponent is playing infallibly, perceived c will move swiftly towards the chosen c_{max}. Given the requirements on $S_c(val, d)$, it may be shown[5] that, as κ increases, R_c progressively loses its ability to differentiate between better and worse moves, that R_c's expectation of state and theoretical value do not improve and that $R_c \to R_0$. Thus, for a given set of observations, an Analyser assuming a greater κ will infer an increasing apparent competence c.

We have recently chosen a fixed $\kappa = 1$, in effect including the immediate move in the line contemplated. We have not tested the effect of different κ on a Predator's choices of move,

[5] The proof is by elementary algebra and in the style of Theorem 3 [6,7].

but assume it is not great. There seems little reason to choose one value of κ over another but the model of the endgame and WILHELM do allow this as a parameter.

4.2 The effect of the initial probability assumption

The usual, neutral, initial stance is a *know nothing* one, assuming that c is uniformly distributed in a conservatively-wide interval $[c_{min}, c_{max}]$. However, it is clear that had BELLE been using the REP model, it could have started game two with its perception of Browne as learned from game one, just as Browne started that game with his revised perception of KQKR. Also, one might have a perception of the competence c likely to be demonstrated by the opponent with the given endgame force – and choose this to be the mid-point of a $[c_{min}, c_{max}]$ range with a normal distribution.

Bayesian theory, see Section 3.2 above, shows that the initial, assumed non-zero probabilities continue to appear explicitly in the calculation of subsequent, inferred probabilities. We therefore note that initial probabilities have some nominal effect on the inferred probabilities but that this effect decays as subsequent experience takes over.

4.3 The effect of the chosen metric

The metric Depth to Conversion (DTC) was chosen because *conversion* is a common intermediate goal: capturing BELLE'S Queen was Browne's objective. The adoption of DTC is however a chessic, domain-specific decision, even if it is an obvious one. Our analysis of the Browne-BELLE games shows that the Predator would never have made a DTC-suboptimal move-choice for Black. It is reasonable to assume that, had DTM(ate) been the chosen metric, it would never have chosen a DTM-suboptimal move. Different metrics often define the same sets of optimal moves but these sets can diverge and even become disjoint as the goals of those metrics approach. Where this occurs, the Predator would choose a different move in its tracking of the Browne-BELLE games.

5 A Markov model of the endgame

Let us suppose that the Preference Function $S_c(val, d)$ is fixed, e.g., as the function defined here with $\kappa = 1$. Given a position P in state s_i, we can calculate the probability of R_c choosing move m to some position P' in state s_j. We may therefore calculate the probability, $T_c(j)$ of moving from position P to state s_j. Averaging this across the endgame over all such positions P in state s_i, we may derive the probability $m_{i,j}$ of a state-transition $s_i \rightarrow s_j$ assuming initial state s_i. The $\{m_{i,j}\}$ define a Markov matrix $\mathbf{M_c} = [m_{i,j}]$ for player R_c. This matrix, and the predictions which may be derived from it, provide a characterisation of the endgame as a whole.

Let us assume that the initial position is 1-0, in state s_i, and that R_c does not concede the win. From the matrix, we may calculate, as shown in Appendix C:

- the probability of R_c (starting in state i) being in state j after m moves,
- the expected depth after m moves,
- the probability of winning from state i in m moves or less,
- the probability of winning from state i in exactly m moves,
- the expected length of win for R_c starting in state I, achieve the win.

These theoretical predictions were computed for KQKR and compared with the results of the extensive experiment described in the next section. Perhaps counter to intuition, there is no minimum capability *c* below which a win is impossible; quite the opposite. Because the win is assumed to be retained, it will eventually be achieved, if only because R_c executes an unlikely optimal move sequence.

Table 1. Statistical Analysis of the 2,000-game experiment.

KQKR: R_{20} - R_∞	**Position 1**	**Position 2**	**Overall**
Min., end-of-game apparent *c*	15.06	14.73	14.73
Max., end-of-game apparent *c*	35.66	40.71	40.71
Mean, end-of-game apparent *c*	21.318	21.620	21.469
St. Dev., end-of-game apparent *c*	3.345	3.695	3.524
St. Dev of the Mean apparent *c*	0.106	0.117	0.079
|Mean *c* - 20|/Stdev_mean	12.43	13.85	18.59
Min. moves, *m* , to conversion	37	37	37
Maximum moves, *m*	395	325	395
Mean moves, *m*	96.88	94.31	95.60
St. Dev., *m*	102.951	102.273	102.587
St. Dev., mean of *m*	3.256	3.234	2.294

6 An experiment with r_{2o}

Echoing Browne-BELLE, a model KQKR match was staged between the fallible attacker R_{20} and the infallible defender R_∞. It was assumed that R_{20} would not concede the win but eventually secure it as theory predicts. The game-specific repetition and 50-move drawing rules were assumed not to be in force. Table 1 summarises the results of this experiment. 1,000 games were played from each of the two maxDTC KQKR positions (DTC = 31) used in the Browne-BELLE match. Games ended with mate or capture of the Rook. The purpose of the experiment was to observe:

- the distribution of the *c* inferred by an Analyser[6] at the end of each game with the assumed probability of c_i set to 1/51 at start of each game
- the distribution of the lengths of the games, and
- the trend in the Analyser's inferred *c*, ignoring game-starts after the first.

The mean game-length of 95.60 and standard deviation of 2.294 show the experiment agreeing closely with the theory. The Markov matrix predicts a mean game-length of 97.20 for *c* = 20 and 83.70 for *c* = 21. Ignoring game starts and ends, the Analyser correctly identifies the capability of R_{20} as 20. Starting afresh from the start of each game, the Analyser shows a mean end-game apparent *c* of 21.50.[7]

[6] using $c_{min} = 0$, $c\delta = 1$ and $c_{max} = 50$ as found adequate in Section 4.1.
[7] Shorter games yield higher end-of-game apparent *c* which are more widely distributed.

7 The KBBKN data

Having checked that experiment and theory were confirming each other, we turned to another classic 5-man endgame, KBBKN [11]. More men implies more positions, greater depths and larger Markov matrices. Calculations were carried out in double-precision arithmetic to ensure that sufficient precision was retained in creating and using the matrices.[8]

Some characteristics of the theoretical predictions are similar to those of the KQKR data; others are different. Again, progress both at the most extreme depths and at shallow depths, seems easier than at the intervening depths where near-optimal moves are plentiful and hardly distinguishable from optimal moves. Again, there is exponential decay, after an initial peak, in the probability of a win in exactly *m* moves. It is clear that KBBKN is more difficult than KQKR as one might expect. In terms of the REP model, a higher capability *c* is required to win KBBKN with a similar efficiency to a KQKR win.

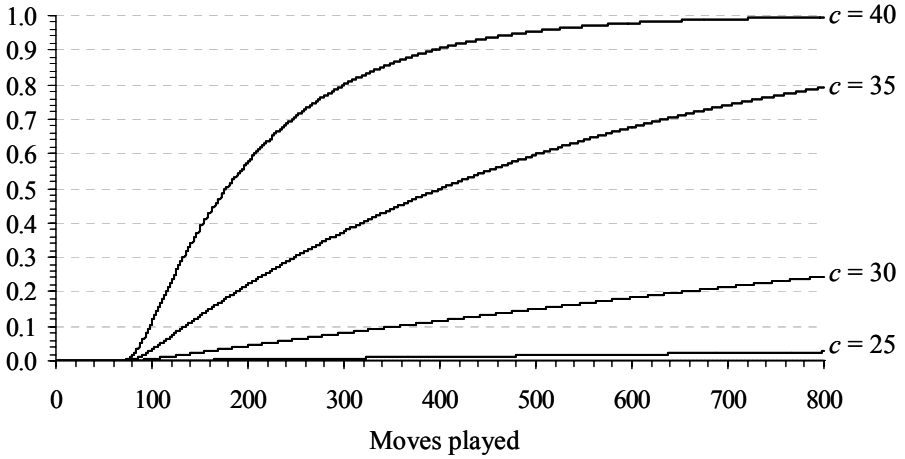

Figure 1. Probability[conversion from maxDTC position in ≤ *m* moves].

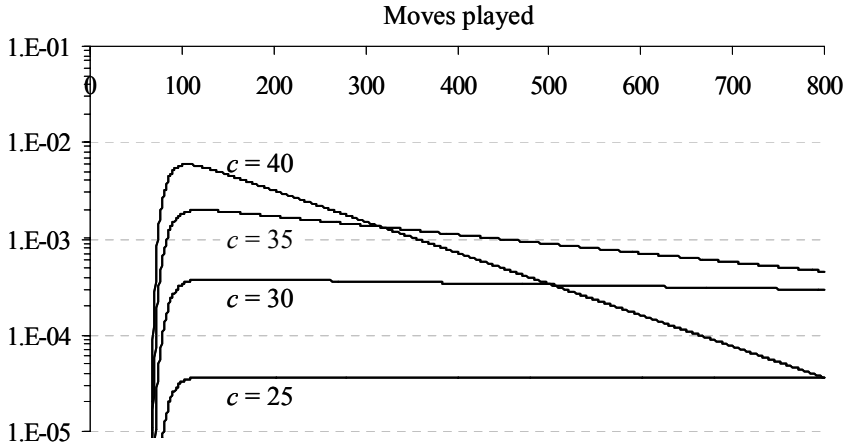

Figure 2. Probability[conversion from maxDTC position in *m* moves].

[8] Matrix $\mathbf{I} - \mathbf{M}_c$ has condition-number $\sigma_1/\sigma_{68} < 10^8$, leaving 7 significant figures accuracy.

7.1 The probability of winning

Figures 1 and 2 show the probability of winning, respectively, in up to and in exactly m moves. The latter probability peaks at a slightly larger number of moves as c is reduced. The games were played without the 50-move rule but the Markov model would, if required, allow us to calculate the probability of winning from depth d on or before move 50, before a possible draw-claim by the opponent. That probability is the probability of being in state 0 after 50 moves, namely the element $\mathbf{M_c}^{50}[d, 0]$ of $\mathbf{M_c}^{50}$.

Figure 3 gives the expected length of an R_c-R_∞ game for each initial depth to maxDTC. Note that for $c = 20$, and starting at depth 31, KQKR games are expected to take 97 moves while KBBKN games average 3,444 moves. Figure 4 gives these probabilities of R_c, winning in 50 moves from any initial depth to maxDTC = 66[9] and for $c = 15, 20, \ldots, 40$.

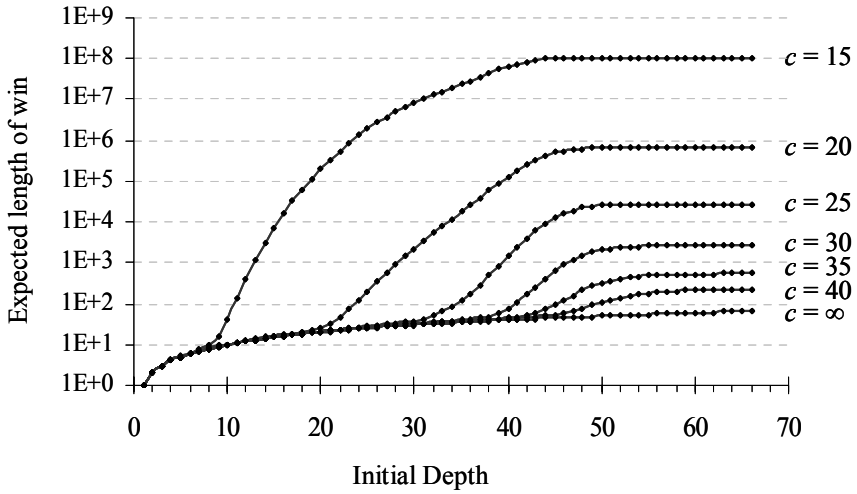

Figure 3. Expected Moves to conversion in a R_c-R_∞ KBBKN game.

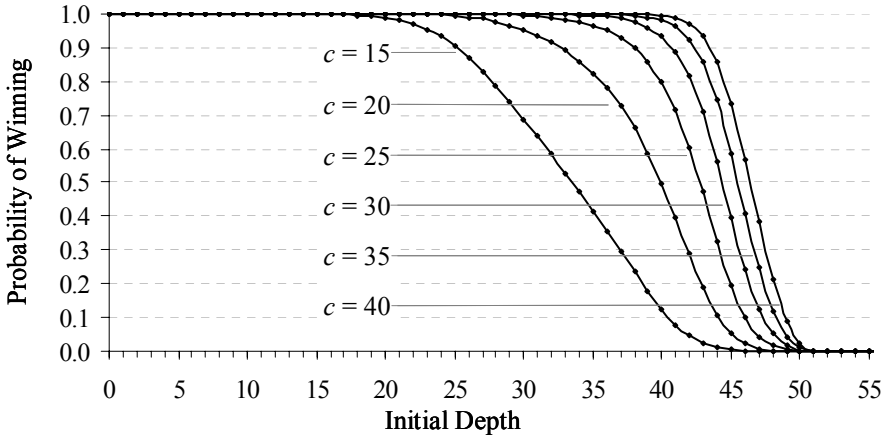

Figure 4. Probability[R_c wins an R_c-R_∞ KQKR game in 50 moves].

[9] This probability is zero of course for initial depths 51-66.

8 Summary

We have examined the utility of a reference model of Fallible Endgame Play by both experiment and theory, using both a comprehensive REP implementation in WILHELM and Markov methods. Various demonstrations have shown opportunities for exploiting the model, and the robustness of decisions based on it. Experimental results have also been compared with the Markov predictions, with which they agree closely.

A comparison of the Markov predictions for KQKR and KBBKN demonstrates some characteristics persisting in the predictions. It also shows that the greater depths of KBBKN, maxDTC = 66, call for greater REP capability c to achieve the same efficacy as in KQKR, maxDTC = 31.

Experiments which remain to be carried out include:

- infallible White attacking fallible Black in a drawn position
  e.g., in KBBKN, KNNKP, KNPKN, KQNKQ, KQPKQ or KRBKR,
- infallible Black pressing for a draw in a lost position
  this requires additional EGT data on draws forced in d moves,
- a more insightful Predator searching more than $2p$ plies ahead, and
- use of the Emulator as a training partner for human players.

The REP model may be extended to other games where EGTs may be computed – to convergent games such as Chinese Chess, Chess Variants, 8×8 checkers and International Draughts. If a search-method can propose what it considers the best few moves in a position, each evaluated on an identical basis and therefore comparable, the concept of a stochastic player may be applied more generally than to just endgames for which perfect information is available.

Acknowledgements

We thank the referees for their feedback on this paper. We thank Walter Browne for his excellent sporting example in facing up to Ken Thompson's silicon beast BELLE in the KQKR match of 1979. Also, we thank those who have generated and/or made available definitive EGT data over the years, especially John Tamplin [12] and Mark Bourzutschky [2] for their DTC EGTs.

References

[1] R. Andrist. http://www.geocities.com/rba_schach2000/. WILHELM download, 2004.
[2] M. Bourzutschky. Private communication of generalized EGT-management code, 2003.
[3] D.M. Breuker, L.V. Allis, H.J. van den Herik, I.S.Herschberg. A Database as a Second., ICCA Journal 15 (1) (1992) 28-39.
[4] W. Feller. An Introduction to Probability Theory and its Applications, Vol. 1. Wiley, 1968. ISBN 0-4712-5708-7.
[5] C.J. Fenner. Computer Chess, News about the North American Computer Chess Championship, The British Chess Magazine 99 (5) (1979) 193-200.
[6] G.M^{c}C Haworth. Reference Fallible Endgame Play, in: J.W.H.M. Uiterwijk (Ed.), Proceedings of the Seventh Computer Olympiad Computer-Games Workshop, Maastricht, Report CS-02-03, 2003.
[7] G.M^{c}C Haworth. Reference Fallible Endgame Play, ICGA Journal 26 (2) (2003) 81-91.

[8] G.M^{c}C Haworth, R.B. Andrist. Model Endgame Analysis, in: H.J. van den Herik, H. Iida, and E.A. Heinz (Eds.), Advances in Computer Games 10, Kluwer Academic Publishers, Norwell, MA, 2003, pp. 65-79. ISBN 1-4020-7709-2.
[9] P.J. Jansen. KQKR: Awareness of a Fallible Opponent, ICCA Journal 15 (3) (1992) 111-131.
[10] D. Levy, M. Newborn. How Computers Play Chess, esp. pp. 144-148, 1991.
[11] A.J. Roycroft. Expert against the Oracle, in: J.E. Hayes, D. Michie and J. Richards (Eds), Machine Intelligence 11, Oxford University Press, Oxford, 1988, pp. 347-373.
[12] J. Tamplin, J. Private communication of some pawnless Nalimov-compatible DTC EGTs, 2001.

Appendix A: Acronyms, Notation and Terms

Analyser	an agent identifying a fallible opponent as an R_c player
c	the competence index of an REP
$c\delta$	the difference between adjacent c_i assumed by the Analyser
c_{max}	the maximum c assumed possible by the Analyser
c_{min}	the minimum c assumed possible by the Analyser
d	the depth (of win or loss) of a position in the chosen metric, e.g. DTC
DTC	Depth to Conversion, i.e. to change of material and/or mate
DTM	Depth to Mate
Emulator	an agent, E_c, choosing moves to best exhibit apparent competence c
Horizon	a search limit, within which R_c will win or not lose if possible
κ	$\kappa > 0$ ensures that $(d + \kappa)^{-c}$ is finite
λ	a scaling factor, matching the probability of loss to that of a draw
L_i	expected length of win (to conversion in winner's moves) from depth i
maxDTC	maximum DTC (depths)
$\boldsymbol{M_c}$	a Markov matrix $[m_{i,j}]$
$m_{i,j}$	the probability, averaged over the endgame, that R_c in state s_i moves to s_j
metric	a measure of the depth of a position, usually in winner's moves
n	the number of different c_i assumed by an Analyser
n_1	$n_1 > n_W$, ensures that draws are less preferable than wins
n_2	$n_2 > n_B$, ensures that draws are more preferable than losses
n_B	the number of 'Black win' states
n_W	the number of 'White win' states
ns	the number of states for a chosen endgame and depth metric
$p(x)\cdot\delta x$	the probability that R_c's $c \in [x, x + \delta x]$
$pc_{0,j}$	the *a priori* (before a move) probability that the unknown c is c_j
$pc_{i,j}$	the probability, inferred after the ith move, that the unknown c is c_j
Player	an R_c, choosing its moves stochastically with Preference Function S_c
Predator	an agent, choosing the best move possible on the basis of an opponent-model
Predictor	an agent predicting the longer term prospects of a result from Markov theory
REP	Reference Endgame Player
R_0	the REP which prefers no move to any other
R_c	an REP of competence c
R_∞	the player which plays metric-optimal moves infallibly
s	endgame state
s_i	(endgame) state i
$S_c(val_s, d_s)$	the Preference Function for REP R_c, a function of destination value and depth
$S_c(val, d)$	a convenient contraction of $S_c(val_s, d_s)$
$S_c(s)$	a more convenient contraction of $S_c(val_s, d_s)$
$T_c(i)$	the probability that R_c moves to state i, s_i
val	the theoretical value of a position, i.e., *win*, *draw* or *loss*
v_m	a weighting that may be given to a move on chessic grounds

Appendix B: Preference Functions

We require that the set $\{R_c\}$ is in fact a linear, ordered spectrum of R_c players such that:

- for R_0, all moves are equally likely,
- 'R_∞' ≡ $\lim_{c\to\infty} R_c$ exists and is the infallible player choosing metric-optimal moves,
- '$R_{-\infty}$' ≡ $\lim_{c\to-\infty} R_c$ exists and is the anti-infallible player choosing anti-optimal moves,
- $c2 > c1 \Rightarrow R_{c2}$'s expectations of successor state, i.e. E[s], are no worse than R_{c1}'s,
- $c2 > c1 \Rightarrow R_{c2}$'s expectations of theoretical value, i.e. E[val_s], are no worse than R_{c1}'s.

The following requirements on $S_c(val, d) \equiv S_c(s)$ are natural ones and sufficient to ensure the above, as proved in [6,7]:

- $S_c(s)$ is finite and positive: no move has zero or infinite preference for finite c,[10]
- $S_0(s)$ is a constant,
- for some $n_1 > n_W$ and $n_2 > n_B$, $S_c(draw) = S_c(win, n_1) = S_c(loss, n_2)$,
- $F_j(c) \equiv S_c(s_{i+1})/ S_c(s_i)$ decreases as c increases: $\lim_{c\to\infty} F_j(c) = 0$ and $\lim_{c\to-\infty} 1/F_j(c) = 0$,
- for $c \neq 0$, $\text{sign}(c)\cdot S_c(s_j)$ decreases (↓) as j increases (↑),
- for $c >$ (<) 0, $W_c(d) = S_c(win, d)/ S_c(win, d+1) \downarrow$ (↑) as $d \uparrow$ and $\lim_{d\to\infty} W_c(d) = 1$,
- for $c >$ (<) 0, $L_c(d) = S_c(loss, d+1)/ S_c(loss, d) \downarrow$ (↑) as $d \uparrow$ and $\lim_{d\to\infty} L_c(d) = 1$.

The net effect is that:

- the spectrum of R_c is centred as required on the random player, R_0,
- the R_c with $c > 0$ prefer better moves to worse moves,
- the R_c demonstrate increasing apparent skill as $c \to \infty$,
- R_c can be arbitrarily close to being the metric-infallible player for finite c
- as $d\to\infty$, R_c discriminates less between a win (or loss) of depth d and one of depth $d+1$.

Appendix C: REP Markov Matrices

After making decisions about the various parameters of the REP model, Markov matrix $\mathbf{M_c} = [m_{i,j}]$ defines for player R_c the average probability, m_{ij}, of R_c moving to state j given that it is in state i. Let us assume that the position is a 1-0 win. Then, if R_c is in fact the infallible defender R_∞, $\mathbf{M_c} \equiv \mathbf{I}$, the identity matrix. This is because depth of win is measured in winners' moves, and therefore losers' moves do not change the depth. Let us assume, as in the experiments, that R_c is a fallible attacker against an infallible defender but that R_c never loses sight of the win. If the initial state-probability vector is $\underline{\mathbf{p}}_0^T$:

- $\underline{\mathbf{p}}_\mathbf{m}^T = \underline{\mathbf{p}}_0^T.\mathbf{M}^\mathbf{m}$ is the state-probability vector after m moves
- $p_{m,j}$ = Pr[being in state j after m moves]
- $\Sigma\, p_{m,j}.d_j$ = E[depth after m moves]
- $p_{m,1}$ = Pr[being in state 1, i.e. having won after m moves]
- $p_{m,1} - p_{m-1,1}$ = Pr[winning in exactly m moves]

Let l_i be the expected length of win from state i. Then $l_1 = 0$ by definition. Otherwise:

$$l_i = 1 + \sum_j m_{i,j}.l_j \Rightarrow -1 = \sum_{j\neq i} m_{i,j}.l_j + (m_{i,i} - 1).l_i \Rightarrow (1 - m_{i,i}).l_i - \sum_{j\neq i} m_{i,j}.l_j = 1$$

Thus the equations $\mathbf{A}.\underline{\mathbf{L}} = \underline{\mathbf{U}}$ solve for $\underline{\mathbf{L}} = \{l_i\}$ where:

$\underline{\mathbf{U}} = (0, 1, \ldots, 1)$ and $\mathbf{A} \equiv \mathbf{I} - \mathbf{M_c}$ except that $A_{1,1} = 1$.

The number of significant figures in computations of l_i depends on the precision of the arithmetic and the condition number of $\mathbf{A}$ which was therefore checked using MATLAB. Condition number is observed to increase as c decreases until, eventually, the l_i for R_c are effectively incalculable in the double-precision arithmetic used. For KBBKN, the condition numbers for $c = 15$ was $4*10^8$ and for $c = 9.95$ was $3*10^{12}$, still yielding significant results.

[10] Hence the requirement that $\kappa > 0$, to accommodate the case of $d = 0$ in $(d + \kappa)^{-c}$.