TABLE OF CONTENTS

TIME TO EVALUATE

We shall supply you with a position. Given that, is it true that the perfect evaluator is the perfection of all Grand Masters or even the grandest of them all, the World Champion? Or could it happen, at least on occasion, that the perfect evaluator was a perfect fool?

While most people are willing to concede that the best evaluation is worthy of mastership, not all evaluations by Grand Masters are masterly. Conversely, the grandest masterly move need not imply a perfect evaluation.

We do not even intend a pun when we state that now you perhaps see your Editors' position in a philosophical sense. But not just a philosophical sense could be intended, we blithely have given you a *position*. An arbitrary one, just any? Or a most definite one? As a testcase for your mastership or folly as the case may require?

Evaluation, specifically as embodied in the evaluation function at least was a solid anchor for conventional computer-chess thoughts only a few years ago. (We note with satisfaction that our discipline is now so mature that the notion of *conventional* thinking makes sense.)

In this issue of the Journal some conventional, almost intuitive notions about evaluation are in for a rough shaking. For a mildly shattering example, consider what Beal and Smith feel duty-bound to communicate.

An evaluation which knows no more about chess than the simplest rules of the game, how to recognize a checkmate and *how to generate a random number* is much, much better than pulling random legal moves out of a hat. What a message for the true believers of yore: do not bother learning about chess, computers are there to generate reliable pseudo-randoms!

Another shock is administered from an unexpected quarter: Iida stemming from Japan, a Grand Master of Shogi, has found partners in the Netherlands to argue plausibly that one's own evaluation function must take second place and, indeed, can often be wilfully and usefully disregarded. This is not Oriental self-effacement, quite the contrary. With utterly practical astuteness, it is argued that knowing the evaluation, wrong or even right, as practised by one's opponent can be used to great advantage to undermine that very opponent. The art of one's own evaluation may well have to yield its habitual primacy to the higher, perhaps martial art of predictively thwarting what the opponent is conjectured to come up with. In particular, there is no mercy for an opponent who is rash enough to have a speedy, precomputed riposte. He is only too likely to have searched less deeply than he should have, which will prove his undoing.

The evaluation function traditionally has been elevated on a high pedestal: it was presumed to comprise all we always wanted to know about chess but were afraid to ask. To *comprise* all? – rather, to compress it all into a few terms. Such compression must surely be one deep calling another. This pedestal, if we are to believe the stated, well-argued conviction of Dr. Morales (Mexico), must be taken several notches down. Chess masters, it is widely believed, base their skills on pattern recognition. If so, let us build a program capable of deriving chess patterns from a few simple instances fed to them and some general-purpose knowledge of chess. The surprising thing in the Morales' study is how pitifully little chess knowledge, how derisively few concepts are needed in order to generate at least some of the much-vaunted patterns. If patterns are paramount, and the capability of employing them as primaries in the evaluation function is the hallmark of a Grand Master, then Morales makes a strong case for an unsuspectedly elementary nature of those patterns. He persuades us, in doing so, to lower our estimate of Grand Masters and their presumed built-in evaluation functions considerably.

While not claiming that his domain of application, the KRK endgame, is a particularly subtle example of the art of pattern exploitation – which it evidently is not – yet, Morales' moral position is that the structure of the Grand Master as a pattern recognizer has been overrated in complexity. Just put in the notion of co-ordinates, of a distance or a Manhattan distance, suitably restricted where necessary, add opposition in its crudest geometrical sense and concepts, such as 'moving closer', 'Kings are on the same side' and 'almost in opposition', roll out with disconcerting ease.

Rather more obliquely, a report by Jonathan Schaeffer in this issue hints that an evaluation function in checkers, however well-exploited, is subject to the law of diminishing returns to the extent that additional plies of evaluation-function-guided search are pointless in the end, regardless of the speed-up of the searching computer. This, as your Editors read it, must be taken as yet another questioning of the evaluation function as the be-all and end-all of chess.

Your Editors are not in the habit of providing ready answers to the questions they raise, nor are they capable of doing so in the present instance. They *do*, however, insist that this is a propitious season to raise a pressing question, a fit moment to ask crucially: is it not time to evaluate the evaluators?

Bob Herschberg
Jaap van den Herik