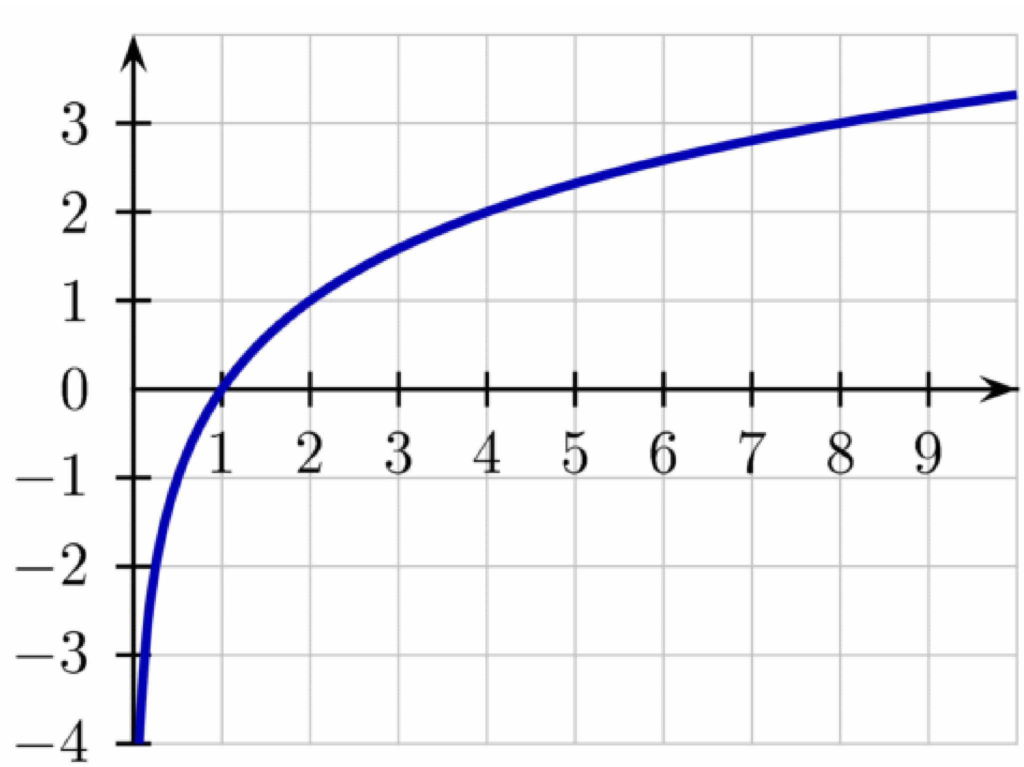
Information – Entropy theory of Artificial Intelligence



The basic theory on which one chess program can be constructed is that there exists a general characteristic of the game of chess, namely the concept of entropy.

This concept has been employed in physics for a long time. In the case of a gas, it is the logarithm of the number of those microscopic states compatible with the macroscopic parameters of the gas.

What does this mean in terms of chess? A common characteristic of every piece is that it could move to certain squares, including by capture. In any given position, therefore, the pieces by the rules of the game possess certain states, only one of which will be realized on the next move. The difference of the logarithm of the numbers of such states for Black and White respectively is the "entropy of the position". The task of the computer is then to increase this value for its own benefit.

Every chess player knows that the more mobility his pieces have and the more constrained are his opponent's, the better his position. For example, checkmate is the best possible state for the attacker, and the chess program playing according to the above principle without the prior notion of checkmate will automatically attempt it if possible.

Entropy is a principle of statistical physics and therefore is only applicable in statistical contexts. The number of microstates of a confined gas is very large and therefore the statistical approach is valid. In chess, however, the number of pieces, a macroscopic parameter, is very small and therefore in this context the "value" of a position cannot be an exact function of entropy. For example, it is possible to checkmate with a total force of a single pawn despite the fact that the opponent has many pieces and various positions available.

Examples of sacrificial combinations further demonstrate this consideration. Therefore we also need specific information about any given position. For example, entropy could be maximized by White giving check, but if the checking piece is then taken the move was a bad one. The logarithm of the number of variations which have been examined in this way gives the amount of information. In the endgame it is rather inaccurate. Because of the small number of pieces the above noted inadequacy of the statistical principle becomes evident and we need to compute much more information to fill the gap.

We can think about the positive logarithmic values as the measure of entropy and the negative logarithmic values as the measure of information.

Shortly speaking:

- The evaluation of any position is based on the entropy + information.
- The entropy is the logarithm of the possible legal moves of the position.
- The information is simply the depth of the search, since it is the logarithm of the exponential growing number of possible positions, $\log e^x = x$.

E = entropy

I = information

D = depth of search

M = legal moves in any position, M_w for white moves and M_b for black moves

$E = \log M_w - \log M_b = \log M$

And since $\log e^x = x$, $I = D$

We get information + entropy, the value V of any position in the search tree of the current chess position:

$V(D,M) = I + E = D + \log M$

This naturally gives better values for deeper search with greater mobility.