The Glicko-2 System for Rating Players in Head-to-Head Competition

Professor Mark E. Glickman
Boston University
July 2000

The Glicko-2 rating system is intended to estimate the skill of player performance in head-to-head competition in a variety of games. In organized gaming, a rating system is useful for players to be able to track their skill strength, and for the organizers to recognize accomplishments of their membership via top player lists. Many online gaming sites already have implemented rating systems, but most of these are based on the system developed by Elo in the early 1960's. This system's fundamental weakness is that players' ratings are treated as if they are estimated equally precisely. No distinction is made between players whose ratings are precise measures of their strength versus imprecise measures. A consequence is that ratings do not stabilize when a player's skill is well estimated. On some online chess servers (internet chess club, for example), it is not uncommon for players' ratings to change by order of magnitudes after a couple hours of gameplay due to variable results. Thus the resulting ratings are very imprecise estimates of players' skill.

The Glicko-2 system, founded on some of the basic principles of the original Glicko system, addresses many of the deficiencies in current rating systems by measuring the uncertainty connected with ratings. Furthermore, the Glicko-2 system recognizes the possibility that players can undergo sudden surges in underlying skill, and that this may be reflected in repeated strong performances. In the Glicko-2 system, when a player has performances that are inconsistent with his/her rating, the rating change will be larger than usual because the system recognizes that an underlying skill change may be occurring. Also, the measure of uncertainty will increase when a series of inconsistent results occurs. This reflects the notion that the system is now less certain about the player's true strength due to the unusual performances. As the players' results become more consistent, the uncertainty measure decreases and the rating begins to stabilize again.

Because online gaming sites have become popular, there is clearly competition for their use. Many gaming sites continue to add new games to attract users, as well as offering various levels of memberships, including memberships that require fees. From online discussions among members of such sites, it is evident that there is great interest in having accurate measures of playing strength. Apart from minor variants of the Elo and Glicko rating systems, there are very few attempts at developing systems that estimate playing strength and rank players in online gaming organizations, as well as gaming organizations that are off-line. The Glicko-2 system, or minor modifications of this system, may have great opportunity for becoming the standard in estimating player strength in organized gaming.