Individual Project Proposal:   
Poker Hand Analyzer

David Kasofsky

# 1. General Information

The goal of this analytic is to extract and present information from online poker hand logs. Our analysis would be restricted to No Limit Texas Hold'em, the most popular form of online poker. A single poker site like PokerStars will easily play billions of hands per year and a professional player may play hundreds of thousands if not millions of hands a year. This data can be used for at least the following purposes:

1. Characterizing Players. If one gathers sufficient data about a specific player, patterns in that player's play may become identifiable and thus exploitable when playing against that player.
2. Identifying Profitable Strategies. By examining many hands, one can identify which strategies tend to be profitable. Some machine learning algorithm could be used here, e.g. probabilistic classification to come up with a strategy like “in this case, fold 20% of the time and bet 80% of the time”. Here we'd definitely like the machine learning algorithm to have access to the player statistics gather in the first example.
3. Evaluating play. By combining statistics about particular players, i.e. the results of the first analysis, with guidelines for profitable play, i.e. the results of the second analysis, one may be to able to evaluate a particular player and identify the strong and weak parts of that player's game.

These analyses would be very interesting to both online poker players and online poker rooms. Poker players would be able to learn and improve their play while poker rooms may be able to develop better products, games, and resources for their players.

See next page for data sources.

# 2. General Data Source Information

|  |  |  |
| --- | --- | --- |
| **Data Sources** | **Description** | **Data Size** |
| Poker Client Hand Histories | Online poker clients record logs of each hand the player plays. These are saved locally on the player's computer. Some players' hand histories may be available online or we could ask players to provide their hand histories. Best if these players are professionals who play lots of hands. | MB-GB |
| Poker Hand Replay Websites | There are websites that let poker players share individual hands, e.g. BoomPlayer.com. We could scrape these sites for hands. An individual hand record is very small but we may be able to gather millions. The problem with this data source is that people would usually only share exceptional hands. | MB-GB |

# 3. Detailed Data Source Information

|  |  |  |
| --- | --- | --- |
| **Data Sources** | **Data Characteristics** | **Data Frequency** |
| Poker Client Hand Histories | A single hand history would be loaded once, but nothing would stop us from gathering more hand histories over time. A good minimum number of hands would be about one million. | One time load, ad hoc additional data |
| Poker Hand Replay Websites | We could continuously scrape these sites for new hands. Hard to say exactly what the rate of new hands being posted is but it wouldn't a giant amount since people don't post all their hands, just the exceptional ones. However exceptional hands may also contain a lot of information. | Realtime, KB-MB / day |