Guess Who

Evan N. Moreira
University of Massachusetts Lowell
One University Avenue, Lowell,
Massachusetts, 01854
Evannigel@gmail.com

Prajwal Marendukonda
University of Massachusetts Lowell
One University Avenue, Lowell,
Massachusetts, 01854
maredukondaprajwal@gmail.com

ABSTRACT
Author Keywords
Decision Tree, ID3, Artificial Intelligence, Twenty Questions, Akinator Web genie

INTRODUCTION
Using Artificial Intelligence in games is nothing new. Games continue to be an integral part of the development of AI theory. Twenty Question and Akinator Web Genie are only two of many games using AI. Both of these games use variations of decision tree learning algorithms.

History of Twenty Questions and Akinator Web Genie
Twenty Questions started as a spoken parlor game, which encourages deductive reasoning and creativity. [5] The game became extremely popular during the late 1940's in the United States.

The rules of the game are as follows:
1. One player is chosen to be the answerer, and chooses an object but does not reveal this to the other players.
2. All other players are questioners.
3. Each questioners takes a turn asking a question which can be answered with a simple “Yes” or “No” or “Maybe”
4. The answerer is not allowed to lie

20Q is a computerized game of Twenty Questions that began as a test in AI developed by Robin Burgener in 1988. [1] 20Q AI asks a series of questions before guessing what the player is thinking and learns on its own with the information relayed back from the players who interact with it. In this version of Twenty Questions, however, the player can answer with “Yes”, “No”, “Unknown”, and “Sometimes”. The 20Q AI uses a decision tree to pick the questions and to guess.

The game suggests that the information required to identify an arbitrary object is at most 20 bits. Mathematically, each question is structured to eliminate about half the objects, and this allows the questioner to distinguish between $2^{20}$ objects. This means the most effective strategy is to ask questions that will split the field of remaining possibilities roughly in half each time.

Akinator, the Web Genie is an Internet game and mobile app based on Twenty Questions that can attempt to determine which character the player is thinking of by asking a series of questions. [2] Akinator was developed by three French programmers in 2007. The questions are asked by a cartoon genie.

Game Logic – Decision Learning Algorithm
Decision tree is a tree which each branch node represents a choice between a number of alternatives, and each leaf node represents a decision [3]. The decision tree starts with a root node and is split into branch nodes until all leaf nodes are represented as seen in Figure 1.

![Figure 1. Visual Representation of a Decision Tree](image)

There are 3 widely used decision tree learning algorithms: ID3, C4.5 and ASSISTANT.

What is Guess Who
Guess Who is simplified version of 20 questions and Akinator Web Genie. Guess Who asks twenty questions and guess a character based of those questions. The database has a limited amount of questions and characters. This limits the possible guess and how well the game guess. The database was found in an open source platform. The home page and question page can be seen in Figure 2 and Figure 3, respectively.
Figure 2. Home Page of Guess Who

If the algorithm does not guess correctly it prompts the user to add the character to the database and related information, this can be seen in Figure 4.

Figure 3. Question Page of Guess Who

If the algorithm does not guess correctly it prompts the user to add the character to the database and related information, this can be seen in Figure 4.

Figure 4. Adding Character Page of Guess Who

ANALYSIS OF RESULTS
The game at first was not great at determining whom the answerer was thinking of. The ID3 Decision tree algorithm learns from past games to determine better answers. In short, the more it plays the better the algorithm gets.

There was not enough time to do the appropriate amount of testing to the efficiency and effectiveness of this algorithm. Testing a game like this take a great deal of time to have a large enough data set to learn from.

The metric that would have been used if there was enough time are as follows:

Metric of success:
1. Graph of how many times the game guesses right over the times play

2. Test the same few characters 30+ amounts of time and show the progress of if the game guess it correctly
3. Add new characters and see if the game learns them and how fast it guess it correctly

DISCUSSION
The decision tree learning algorithm used in this project was greatly modeled after the ID3. ID3 decision tree learning is method for approximating discrete-valued target functions in which the learned function is represented by a decision tree. ID3 is a simple decision tree learning algorithm developed by Ross Quinlan (1983). The decision tree is constructed by employing a top-down, greedy search through the given sets to test each attribute at every tree node, shown in figure 2.

To find an optimal way to classify a learning set, the questions asked need to be first minimized, which minimizing the depth of the tree. Thus a function called the entropy function. This function updates the decision tree with information gain from playing, hence why the more the game is played the better it gets. The entropy function used in this algorithm is as follows:

\[ H_b(p) = -\log_2(p/1-p) \]

What the code actually does,

The decision tree is setup in the database of the game. After each response, the decision tree is cut in half and entropy for each possible character is updated. After twenty questions, the character with the highest entropy value is chosen as the best guess. The entropy value updates the database, so for the next game to be played the correlation between questions and character are better understood, whether it is negative (no) or positive (yes).
CONCLUSION

Decision tree learning algorithms are extremely useful. They have been used in expert systems in capturing knowledge. Even though our algorithm did not necessarily show this well, over time it would become better. Some key points about using decision trees are to make sure there is a large enough database and a large enough playlog for the algorithm to learn from. We plan to add learning features to our algorithm to make it more accurate and interactive.

ACKNOWLEDGMENTS

The work described in this paper was conducted as part of a Fall 2014 Artificial Intelligence course, taught in the Computer Science department of the University of Massachusetts Lowell by Prof. Fred Martin. Evan had worked on mining the data through various sources for us to test and train the application and designed few html pages and GUI. Prajwal has worked on developing the core algorithm and collaborated in design of web interface.

REFERENCES


