

MuSweeper: An Extensive Game for Collecting Mutual Exclusions

Tao-Hsuan Chang and Cheng-wei Chan and Jane Yung-jen Hsu

Department of Computer Science and Information Engineering
National Taiwan University

doudi.tw@gmail.com, gattathree@gmail.com, yjhsu@csie.ntu.edu.tw

Abstract

Mutual exclusions provide useful information for learning classes of concepts. We designed MuSweeper as a MineSweeper-like game to collect mutual exclusions from web users. Using the mechanism of an extensive game with Imperfect information, our experiments showed MuSweeper to collect mutual exclusions with high precision and efficiency.

Introduction

Mutual exclusions of categories provide a solution to “semantic drift”, the error accumulation in bootstrap learning approaches. In (Carlson et al. 2010), mutual exclusions are utilized as constraints to guide the Never Ending Learning Loop. However, the information was manually specified in NELL. In this research, we designed *MuSweeper*, a MineSweeper-like game to collect mutual exclusions among categories using category-instance assertions collected by GWAPs (von Ahn 2006), e.g. Virtual Pets (Kuo et al. 2009) and Verbosity (von Ahn, Kedia, and Blum 2006). This paper presents the game design, mechanism, and experiments to show that *MuSweeper* is an effective GWAP for collecting mutual exclusions with high precision and efficiency.

Game Design

MuSweeper is a two-player GWAP for collecting mutual exclusions among multiple concepts. Two concepts are considered mutually exclusive if no instance belongs to both concepts. For example, “food” and “furniture” are mutually exclusive, but “food” and “plant” are not (e.g. “melon” is an instance of both food and plant). In Figure 1, two players *bani* (a) and *nameofevil* (b) take turns selecting an *answer concept* (d) in response to the *query concept* (c). When one of the players selects a *mine*, the other player wins. A player may click the *PASS* button (e) if he/she believes all remaining concepts are mines. The other player may win the game by selecting a safe concept. If both players pass the game in consecutive turns, the game is a tie.

Initially, the “*mines*” in MuSweeper are defined by a given ground truth set on mutual exclusions from category-instance assertions in Chinese (from Virtual Pets) and in

Copyright © 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

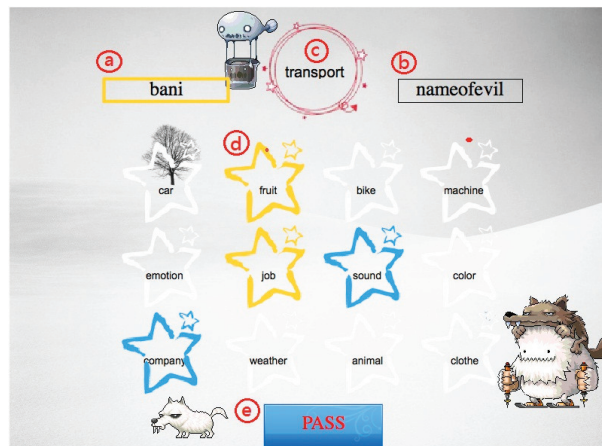


Figure 1: A sample screen of MuSweeper

English (from Verbosity). For example, players ask/answer questions in the form of “__ is a __” to teach commonsense to their pets for *knowledge points* in Virtual Pets.

At the beginning of a game, the purpose of the game, the game rules together with examples/counter-examples of mutually exclusive concepts are presented to the players. To ensure at least one mine exists in each game, MuSweeper chooses randomly some concepts that are known to have common instances from the given ground truth set. Selections by players are collected as new assertions about mutual exclusion among the concepts, aggregated according to the frequency and order of concept selection by the winner of each game. MuSweeper is available in both English and Chinese to any online player.

Mechanism

MuSweeper is modeled as an extensive form game with imperfect information. Players are assumed to share commonsense knowledge on mutual exclusions in terms of an ordering of the probability of two concepts being mutually exclusive. For example, Figure 2 shows a sample ordering of all answer concepts according to their individual probability that it is mutually exclusive with the query concept. A concept on the left-hand side is more likely a mine than the ones on the right-hand side. Given that MuSweeper is a zero-sum

game, rational players should pick the safest choice at any given time to maximize their chances of winning. As a result, the early selections by both players have high probability to be mutually exclusive with the query concepts, so they are collected as labels of mutual exclusions. In contrast, mines are labeled as *not* mutually exclusive with the query concept.

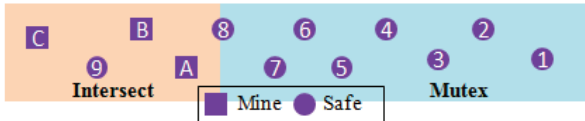


Figure 2: Model of players' selection.

To avoid losing the game, rational players will *not* pass the turn as long as there are *safe* choices, concepts in the blue area in Figure 2, that is, concepts with high enough probability of being mutual exclusions with the query. When no safe choice is available, a rational player will choose to pass. The equilibrium is to pass when there is no choice with *probability* > 0.5 .

To ensure high precision, only the first four selections by the winner player are collected as mutual exclusion assertions in every round. MuSweeper also enjoys the following features:

- **No collusion.** There is no incentive for players to collude in a two-player zero-sum game
- **Efficient.** The game is timed to encourage concept selection at a fast pace.
- **No ambiguity.** There is no free text inputs by the players in the game.

MuSweeper is not designed to collect new concepts; it utilizes a commonsense semantic network that represents hierarchy of concepts to generate mines. More knowledge in the semantic network can help improve the precision of the results by MuSweeper.

Experiments

Experiments were conducted to evaluate the performance of MuSweeper using concepts from the ConceptNet built by Virtual Pets (Kuo et al. 2009), which contains 25 categories with 4 to 20 instances and an average of 11.6 instances in each category. We invited 63 CS-major freshmen locally, and recruited another 44 online users by advertising via a bulletin board system with more than 1.5 million registered users. For 3 weeks, 107 players played MuSweeper against a built-in bot, and the human players won 6,983 rounds out of 8859 games. Each player played 82 rounds on the average, with 63 players played above the average.

Built-in Bot The built-in bot implements random concept selection according to a heuristic probability function that satisfies the following requirements:

- Bot should be undefeated in the first two rounds.

- The probability of losing is positively correlated with the round number.
- Bot will pass only if the human player is expected to pass in the next round.

Precision The precision of the assertions collected by MuSweeper were evaluated based on questionnaires completed by 11 CS-major graduate students. Participants were requested to vote if two given concepts are mutually exclusive or not. A total of 496 pairs were voted true by 8 or more people.

Based on the game log, 290 mutual exclusions were predicted by MuSweeper, only 8 assertions are different from the questionnaire results, giving MuSweeper a precision of 97%. We also found that four of the disagreements are noises, which can be filtered with increasing playing rounds.

Efficiency From the game log, the average time used for each concept selection is 7.203 seconds. An hour of continuous play yields a maximum of 500 mutual exclusions. MuSweeper collected 290 mutual exclusions out of the 496 pairs. The collection efficiency is over 58.5% in 3 weeks.

Conclusion

The experiments showed that MuSweeper is has high precision and efficiency as a valid component for collecting mutual exclusions.

Acknowledgments

This research was supported by NSC 99-2221-E-002-139-MY3 from the National Science Council of Taiwan.

References

- Carlson, A.; Betteridge, J.; Kisiel, B.; Settles, B.; Hruschka Jr., E. R.; and Mitchell, T. M. 2010. Toward an architecture for never-ending language learning. In *Twenty-Fourth Conference on Artificial Intelligence (AAAI10)*.
- Kuo, Y. L.; Lee, J. C.; Chiang, K. Y.; Wang, R.; Shen, E.; Chan, C. W.; and Hsu, J. Y.-j. 2009. Community-based game design: experiments on social games for commonsense data collection. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*.
- von Ahn, L.; Kedia, M.; and Blum, M. 2006. Verbosity: A game for collecting common-sense knowledge. In *ACM Conference on Human Factors in Computing Systems (CHI Notes)*, 75–78.
- von Ahn, L. 2006. Games with a purpose. *Computer* 39(6):92–94.