Sindh Univ. Res. Jour. (Sci. Ser.) Vol.44 (1) 85-90 (2012)

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

Computerized Scrabble Word Game in National and Regional Languages of Pakistan

A.A. SHAH, L.D. DHOMEJA, I. A. ISMAILI AND A. I. AMJAD

Institute of Information and communication technology, University of Sindh, Pakistan

*Corresponding author: e.mail: azhar.shah,lachhman,imdad.ismaili)@usindh.edu.pk. Cell.No.03223077286

Received 16th August 2011and Revised 18th December 2012

Abstract: Abstract: In this paper we propose an algorithm that utilizes the artificial intelligence based techniques including heuristic search and natural language processing to enable the single user to play the game at any time with the computer without waiting for the availability of other player. The novelty of the algorithm lies in the fact that besides English, it also provides support for the vocabulary of the national and regional languages of Pakistan including Urdu, Sindhi and Pashto. In order to make the game more educational in terms of the learning of these languages, speech synthesis based pronunciation .facility is also provided. Thus, the implementation of this algorithm would facilitate any one in general and students in particular to build a strong vocabulary of these languages with just the fun of playing. This would provide a direct contribution to the education and literacy along with entertainment.

Keywords: AI, Heuristic Search, Speech Synthesis, NLP, Tiles.

1. INTRODUCTION

Information and communication technology (ICT) have long been used to aid and enhance the learning skills at almost every level of education. The integration of ICT with the mainstream education has been going under the headings of 'edutainment' and 'play-and-learn'. The integration of play and learning is usually achieved by means of developing computerized educational games. Scrabble is one of these educational games that have become much popular around the world. Besides providing the fun, Scrabble also helps the players to enhance their language learning skills specially spelling and vocabulary. Normally, this game is played by two or four players but its computerized version allows even only one player to play the game (against the computer; the other player). The computerization of this game is based on the use of some heuristic search methods to cope with the imperfect information based specifications of Scrabble (Appel, 1988 Cosma, 1983, Gordan, 1994, Shapario, 1979 Turcan, 1982). The popularity of the game is also evident from the fact that so far it has been implemented in as many as 29 different languages of the world (Horton, 2005). The implementation in each of these languages is the result of the governmental, institutional, and organizational efforts of the respective nations and countries. In Pakistan, so far, little has been done to support the computing in our national (Urdu) and sub-national languages e.g., Sindhi, Pashto, Punjabi, Balochi etc. This is mainly because our educational system gives undue importance to a foreign language (English) while ignoring the languages of our own. Though there are some efforts at individual and institutional level for the promotion of computing in regional languages (e.g., Center for Research in Urdu Language Processing: CRULP, at FAST-NU, Lahore); however, the outcome in this regard is still in its infancy. In order to help foster and enrich the trends for the computerization of local languages, this article proposes an algorithm that utilizes the artificial intelligence based techniques including heuristic search and natural language processing to enable the single user to play the game at any time with the computer without waiting for the availability of other player using the vocabulary of Urdu, Sindhi, and Pashto languages. The developed system also provides speech synthesis based pronunciation facility to help foster the learning of these languages (Shah, 2004). The implementation of this algorithm provides facility to any one in general and students in particular to build a strong vocabulary of our national and regional languages with just the fun of playing. This

would provide a direct contribution to the education and literacy along with entertainment. The rest of the article is organized as follows: section 2 presents a brief description of the AI techniques used for the development of this system; section 3 presents the design and implementation of the algorithms; section 4 presents the results and discussions while section 5 concludes this work.

2. MATERIAL AND METHODS

Though Scrabble is so popular that almost every literate person would be familiar with its name; however, in order to help the reader to understand our algorithms properly and also to make this article self-contained, some details regarding the Scrabble game and its computerization are presented in the following sub-sections.

A. Overview of the Scrabble game

Scrabble was introduced by Alfred Butts, an American-architect in 1938 under the name of "Criss-Crosswords". In 1948, Butts sold this game to James Brunot, an American game manufacturer, who renamed it to 'Scrabble' (meaning 'to scratch frantically'). Within the period of a few years, Scrabble became so much popular that Brunot could not meet the demand for supply for himself and in 1952 he sold the game to a bigger manufacturing company named Selchow and Righter. Since this time, the manufacturing rights of the game have shifted from one company to another and currently it is trademarked by Hasbro, Inc. in the US and Mattel, Inc. elsewhere in the world.

Traditionally, Scrabble is played between two players; however, in the case of tournament games it could also be played between two teams. In this case the members of each team collaborate on a single rack. The game is played on a board marked with a 15-by-15 grid as shown in Fig.1. Each cell in the grid (aka square) would accommodate a single letter tile. Some squares are designated as 'premium'

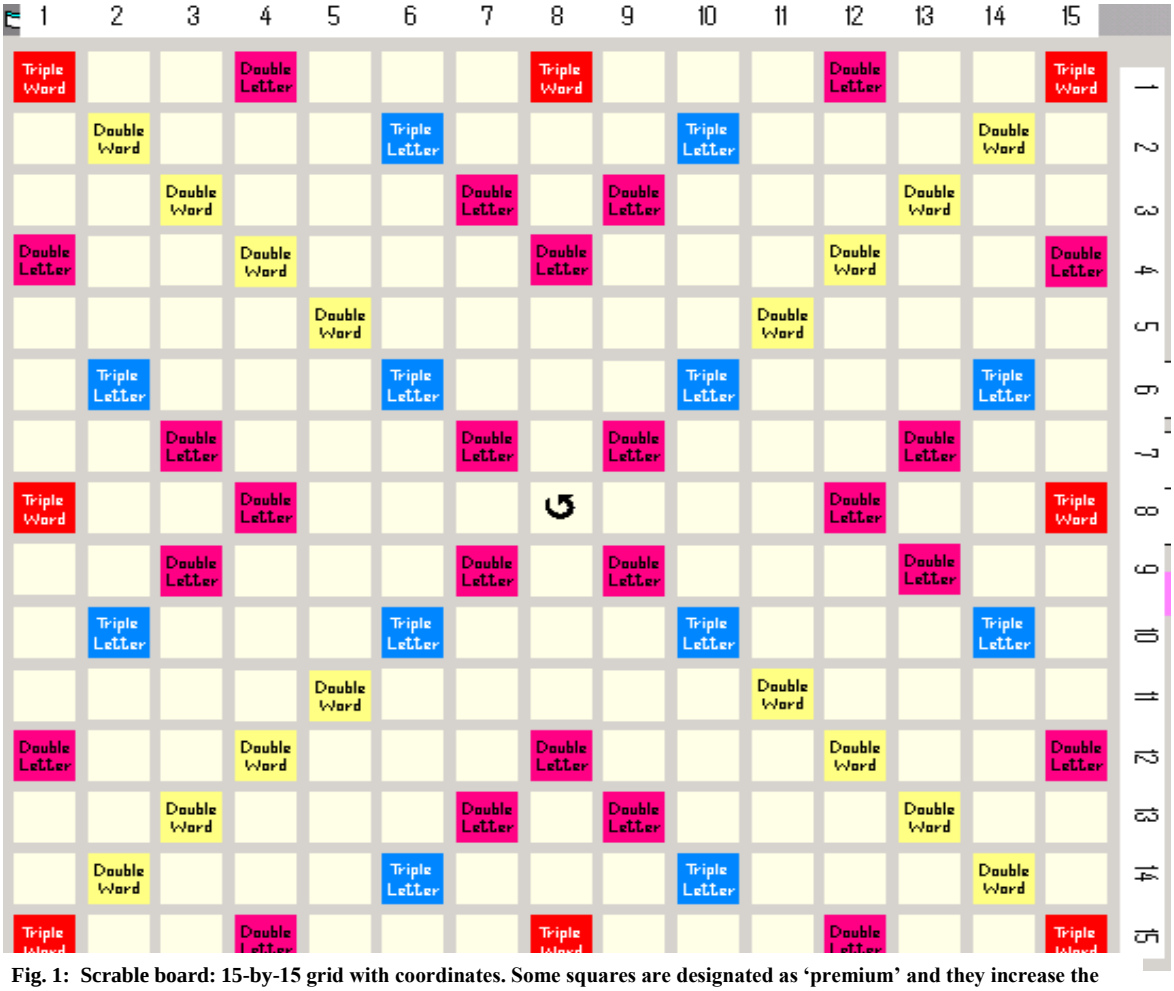


Fig. 1: Scrable board: 15-by-15 grid with coordinates. Some squares are designated as 'premium' and they increase the score of a tile by doubling or tripling the letter score or the word score.

squares and they increase the score of a tile by doubling or tripling either the letter score or the word score. The total number of tiles varies from one language to another. In the case of English, the game consists of 100 tiles. Two of the tiles are left blank (which are provided as an option to be used as a substitute for any letter) while ninety eight of these are marked with a letter and its corresponding scoring point that ranges from 0 to 10. The scoring scheme is based on the frequency of the letter in the writing system i.e., the highest the frequency, the lowest the score (e.g., E and O have the score of 1) and the lowest the frequency, the highest the score (e.g., Q and Z have the score of 10).

B. Overview of the heuristic techniques used for the Scrabble

In Scrabble, the move selection strategy decides which of the move would be best to be taken from a set of all possible moves. The strategy uses a criteria function to evaluate the relative contribution/strength of each move towards achieving the final goal. It should be noted that in terms of Scrabble, the simple greedy heuristic, in which the highest-scoring move is selected in each turn, might not result in providing the best path to victory. This is because of the imperfect information scenarios involved in Scrabble e.g., uncertainty regarding the tiles to be drawn in the next turn as well as the hidden nature of the details of the tiles held by the opponent. This situation is very different from other games including chess and checkers, in which the state of the opponent is perfectly known in advance. Therefore, Scrabble uses the Rack Leave Heuristic as a standard component of its strategy. The Rack Leave Heuristic considers the details of the tiles left on one's rack once a particular move has been made. This consideration enables the strategy to select the tiles which could form high scoring words as well as retain the tiles having high-scoring letters which are expected to be useful at later positions for making even high-scoring words than they could do currently. This sort of approach is achieved by implementing a weighted move selection mechanism based on evaluation of the rack leave strategy with a simple heuristic. Some of the popular computer based implementations of Scrabble for English language include Maven (Sheppard, 2002) and Quackle (Katz-Brown, 2006). Maven is the commercial implementation while Quackle is the open-source alternative. These computer based programs are so intelligent that could even beat their counterpart human experts of high ranking (Read, 2007).

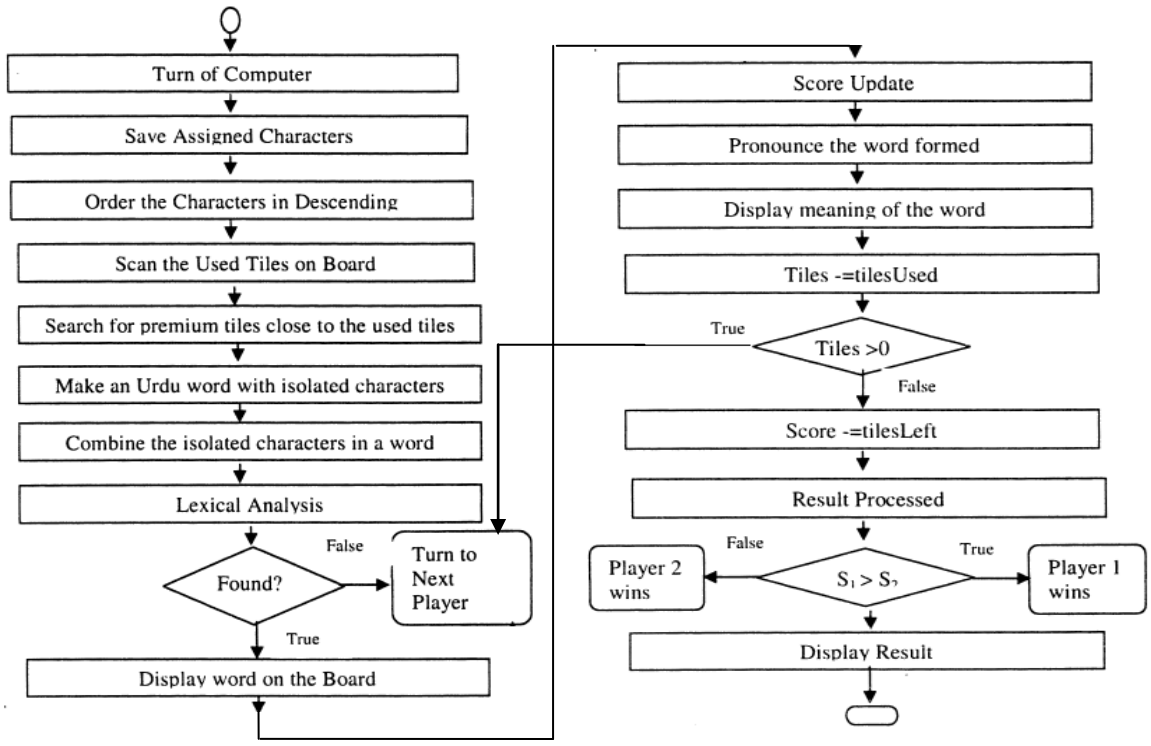


Fig. 2: Algorithm for the turn of computer with Urdu language tiles

2. DESIGN AND IMPLEMENTATION OF THE ALGORITHMS:

The overall algorithm is divided into two parts for each English, Urdu, Sindhi and other regional languages. First algorithm is associated with the turn of computer and second algorithm is associated with the turn of player (user). Here as an illustration we discuss these two algorithms in terms of Urdu language. The design of algorithms for other regional languages would be identical.

C. Algorithm for the turn of computer with Urdu language tiles

This algorithm uses Heuristics search techniques to minimize the search space to make a word with highest score. Fig.2 illustrates the steps used in this algorithm. The algorithm starts with taking care of the available characters by saving and sorting them in descending order. It then scans which squares have been used on the board and searches for premium squares within the proximity of the used squares to make a high scoring Urdu word with isolated characters. The algorithm then uses the additional forms of the Urdu characters e.g., initial, medial, and final to create a properly written word as per recursive nature of Urdu and compares it against the standard Urdu dictionary. If the newly formed word exists in the dictionary, then the word is

displayed on the board along with its meaning and text to speech entry for pronunciation. The score is updated and the game continued for the next word in case if sufficient unused tiles are available; otherwise, the scores of the computer and the human player (user) are compared and the one having high score is declared as the winner of the game.

D. Algorithm for the turn of Player (User) with Urdu language tiles

Fig.3 illustrates the steps used in this algorithm. This algorithm reads the characters used by the player to make the Urdu word and checks if the newly formed word is available in the Urdu dictionary. If the word is not found the game is turned to the computer (the other player that is based on the algorithm discussed above). If the newly formed word exists in the dictionary then its meaning and pronunciation are displayed on the board and the scores are updated. It then checks if the tiles are available and continues the game, otherwise the results are processed and the winner is declared.

3. RESULTS AND DISCUSSIONS

E. Code Composer Studio IDE

The designed algorithms have been tested through the implementation in Microsoft Visual Basic for English, Urdu and other regional languages of Pakistan including Pashto and Sindhi. Snapshots of the GUI interfaces for each language are illustrated in **Fig.4 (a and b).** The operation and environment of the game in each language is satisfactory, convenient and user-friendly.

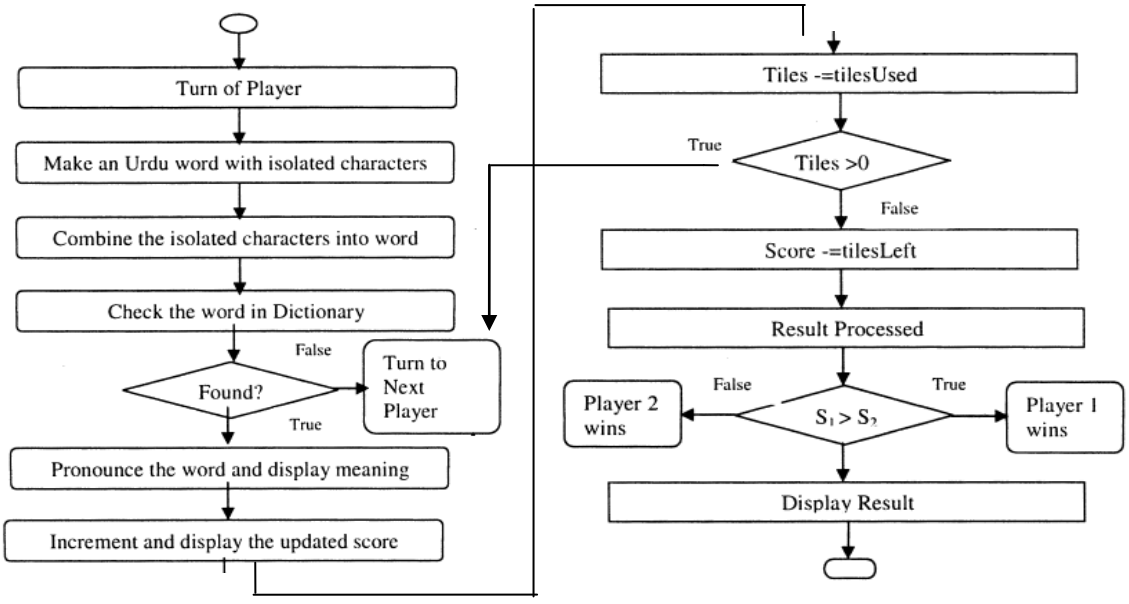
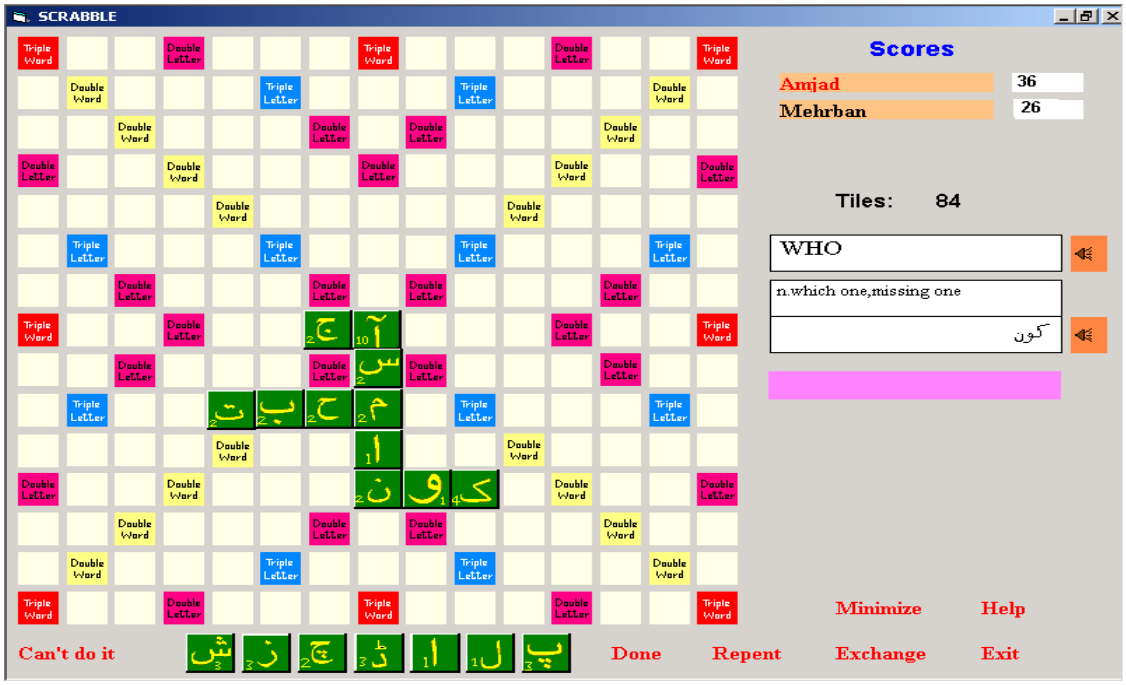
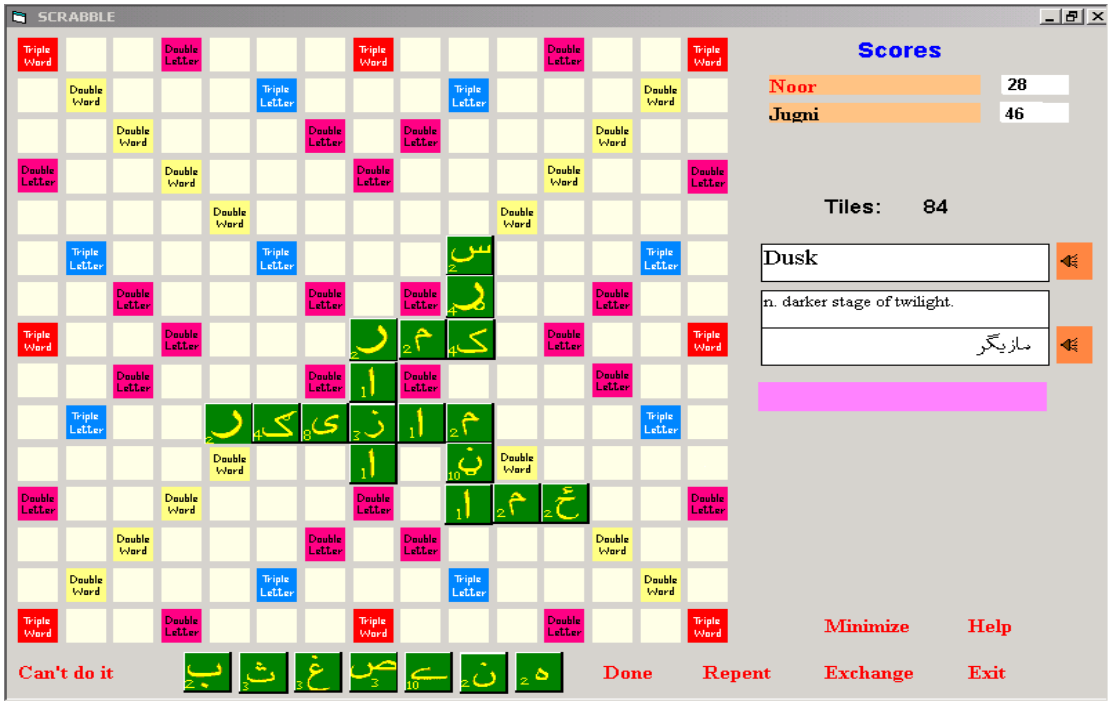


Fig. 3: Algorithm for the turn of Player with Urdu language tiles



(a) Urdu based scrabble GUI



(b) Pashto based scrabble GUI

Fig. 4: Illustrations of the multi-lingual educational game

4. CONCLUSION AND FUTURE DIRECTIONS

A novel algorithm for the adaptation of Scrabble to be played with the alphabet of national (Urdu) and sub-national (Sindhi, Pashto etc) languages of Pakistan has been proposed and implemented in Microsoft Visual Basic. The algorithm is also coupled with text-to-speech conversion system that enables the words to be pronounced correctly. The implementation of this system is yet another step to enrich the computing in our national and regional languages. The system presented in the article could be extended in future by studying the frequency of letters of letters in Urdu and other languages and changing the score scheme that has presently been used as that of English. This will further improve the game and make it more close to its English counterpart.

5. ACKNOWLEDGEMENTS

An earlier version of this article was presented at ICET05, Islamabad.

REFERENCES:

Appel A.W. and G. J. Jacobson (1988) The World's Fastest Scrabble Program. Commun. ACM.31 **(5):** 55Pp.

Cosma J. and D. Jackson (1983) Introducing Monty Plays Scrabble. Scrabble Players News, 710Pp.

Cormen T. H., C. E. Leiserson and R. L. Rivest (2001) Introduction to Algorithms. 5th Edition: 11-16.

Gordon S. (1994) A faster Scrabble Move Generation Algorithm. Software Practice and Experience. 24 **(2):** 23-32.

Shaprio S. C. (1979) A Scrabble Cross word Game Playing Program. Proceedings of the Sixth, IJCAI.

Turcan P. (1982) A Competitive Scrabble Program. SIGArt Newsletter. **(80):** 104-109.

Horton, N. (2005) Welsh is the new word in Scrabble. URL:

http://news.bbc.co.uk/2/hi/uk_news/wales/4308262. stm: Retrieved: May 12, 2011.

Shah A.A, A.W Ansari and L. D. Dhomeja (2004) Bilingual Text to Speech Synthesis System for Urdu and Sindhi. Proceedings of IEEE. ACM and SZABIST National Conference on Emerging Technologies NCET2004, SZABIST, Karachi, Pakistan.

Katz-Brown J. and J.Laughlin (2006) Open-Source Quackle Crossword Game Software Released to Public. URL: http://www.quackle.org/ Retrieved: May 15, 2011

Read B. (2007) A Computer Program Wins Its First Scrabble Tournament. URL: http://chronicle.com/blogs/wiredcampus/a-computer-program-wins-its-first-scrabble-tournament/2800 Retrieved: May 29, 2011.