

Sztuczna inteligencja grająca w Scrabble

Koncepcja implementacji - dane, struktury danych, algorytmy

Jakub Turek

19 grudnia 2013

Streszczenie

Celem artykułu jest opisanie zbioru koncepcji, które posłużą do implementacji algorytmu sztucznej inteligencji grającego w grę Scrabble w języku polskim. Artykuł analizuje i porównuje dane zawarte w dwóch głównych słownikach wyrazów do gier dla języka polskiego, przedstawia dane statystyczne ułatwiające wprowadzanie heurystyk do algorytmu, a także opisuje metody niezbędne do wyznaczania wszystkich możliwych kombinacji ruchów w danej turze. Autor omawia również podział rozgrywki na fazy gry i przybliża podejście, które pozwala uzyskiwać najlepsze wyniki dla każdej fazy gry.

Wstęp

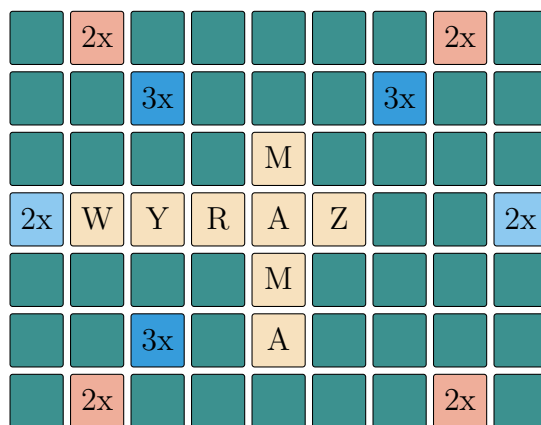
Scrabble to „gra słowna polegająca na układaniu na określonej planszy wyrazów z losowanych liter”¹. Jest to bardzo ogólna definicja, którą należy uściślić. Scrabble jest grą przeznaczoną dla 2-4 osób. Akcesoriami do gry są: kwadratowa plansza o stałym rozmiarze 15 × 15, torebka wypełniona płytkami, na których nadrukowane są litery oraz ich wartości punktowe, a także stojaki, na których gracze umieszczają płytki, którymi w danej chwili dysponują.

¹Wielki słownik ortograficzny - PWN 2003, 2006, 2008 - E. Polański

Gra rozgrywana jest w turach. Zadaniem graczy jest układanie wyrazów na planszy, w taki sposób aby tworzyły one poprawne słowa w języku, w którym prowadzona jest rozgrywka, w układzie krzyżówkowym. Układ krzyżówkowy został przedstawiony na rysunkach 1 oraz 2:

Rysunek 1 Pokazuje sytuację początkową obrazującą pewien moment rozgrywki.

Rysunek 2 Pokazuje poprawny ruch zawodnika, który powoduje powstanie więcej niż jednego słowa. Wszystkie wyrazy utworzone przez jeden ruch muszą być poprawne. W podanym przykładzie słowa „za” i „masz” są poprawne.

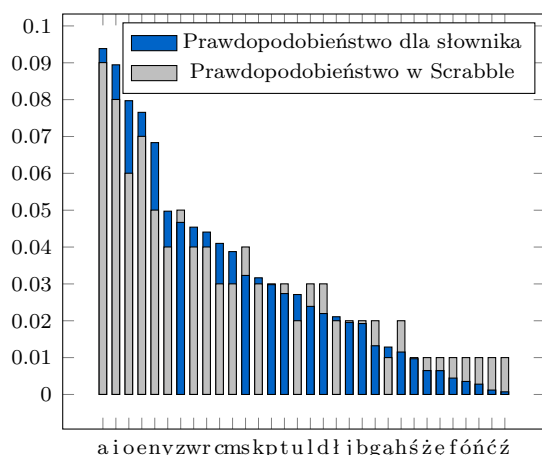


Rysunek 1: Fragment planszy. Gracze ułożyli kolejno słowa: „wyraz” oraz „mama”.

udostępnia listy wszystkich wyrazów zawartych w słowniku. Program pozwala wyłącznie sprawdzić, czy dane słowo jest legalne, bądź nie. Słownik alternatywny dostępny jest w Internecie również w postaci listy wszystkich dopuszczalnych wyrazów zebranych w pliku tekstowym. Pozwala to na przeprowadzenie analiz statystycznych. Stąd w dalszej części artykułu autor rozważa wyłącznie Słownik alternatywny.

Analiza statystyczna

W celu wyznaczenia heurystyk mogących wspomóc sztuczną inteligencję grającą w Scrabble, autor przeanalizował dane zawarte w Słowniku alternatywnym. Pierwszą analizowaną informacją jest częstotliwość występowania poszczególnych liter w słowniku. Wynika to z faktu, że zasady gry Scrabble są oparte na rozkładzie prawdopodobieństwa występowania liter w danym języku. Możliwe jest, że badanie statystyczne słownika przyniesie informacje o literach, których warto używać w pierwszej kolejności.

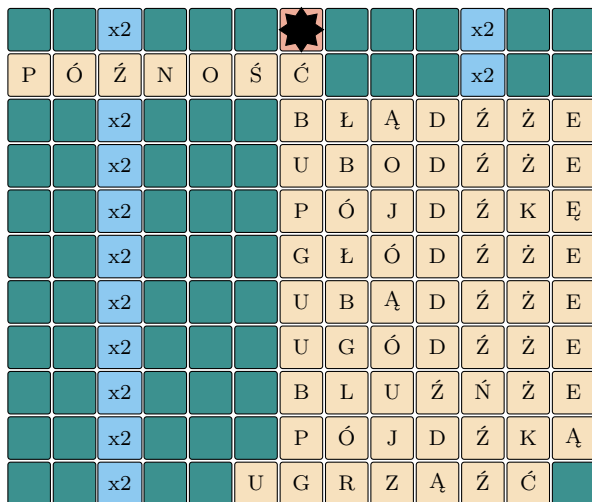


Rysunek 3: Wykres ilustrujący zależność pomiędzy wystąpieniem danej litery w wyrazach słownika, a prawdopodobieństwem wylosowania płytki z tą literą w grze Scrabble.

Wykres przedstawiony na rysunku 3 dostarcza nam informacji, że autorzy polskich liter w grze Scrabble nie oszacowali poprawnie częstotliwości występowania wszystkich liter. Przykładowo, wbrew zasadom gry, litera **o** w rzeczywistości występuje częściej niż litera **e**. Wynika z tego, że porównując dwa podobnie punktowane zagrania, warto wybrać to, które wykorzystuje więcej samogłosek **e** niż **o**. Podobną zależność można stwierdzić w przypadku litery **h**, której częstotliwość występowania została przez autorów zasad znacząco zawyżona. Korzystnie jest wybierać zagrania, które powodują wyłożenie płytek z literą **h**.

Interesującą statystyką są najlepsze otwarcia (czyli pierwsze zagrania w grze). Pozwalają one, przy wylosowaniu szczęśliwej kombinacji płytek, zagrać w najbardziej opłacalny sposób (na początku rozgrywki nie ma ryzyka pozostawienia przeciwnikowi „otwartej gry”). Statystyka ta jest szczególnie przydatna dla ludzi, chociaż również sztuczna inteligencja może wykorzystywać tego typu informacje do pominięcia zbędnych obliczeń. Najlepsze otwarcia dla słownika alternatywnego zebrano na rysunku 4.

Dokonując statycznej analizy słownika, warto wziąć pod uwagę najbardziej opłacalne kombinacje liter. W trakcie rozgrywki gracze dążą do układania wyrazów, które wymagają użycia wszystkich siedmiu płytek znajdujących się na stojaku. Zagranie takie jest bowiem punktowane dodatkową premią w wysokości 50 punktów. Stąd przez najbardziej opłacalną kombinację liter należy rozumieć taką zawartość stojaka, która pozwala ułożyć jak najwięcej wyrazów siedmioliterowych. Duża liczba kombinacji powoduje, że ułożony wyraz uda się wpasować w istniejące na planszy ułożenie płytek z większym prawdopodobieństwem. Autor przeanalizował najlepsze kombinacje liter dla Słownika alternatywnego. Wyniki zebrane są w tabeli 5.



Rysunek 4: Najlepsze otwarcia dla Słownika alternatywnego. Słowo *późność* jest warte 126 punktów, pozostałe po 124 punkty.

Najbardziej opłacalne kombinacje liter można wykorzystać do strategicznego planowania zagrań. Wykonując poszczególne ruchy korzystnie jest dążyć do uzyskania jednego z przedstawionych układów liter na stojaku. Można tego dokonać poprzez usuwanie płytek, które nie zostały uwzględnione w tych kombinacjach, jak również usuwanie płytek zduplikowanych (częstą sytuacją jest posiadanie kilku płytek z tą samą samogłoską).

Litery	Liczba możliwych do ułożenia wyrazów
E, I, K, L, N, O, W	12 słów
A, E, I, K, P, R, S	12 słów
A, E, I, K, L, N, P	12 słów
A, E, K, N, R, T, Y	11 słów
A, I, K, M, O, P, S	11 słów
A, I, K, M, O, R, W	11 słów
A, A, I, K, L, M, S	10 słów
A, I, K, M, O, S, T	10 słów
A, I, K, L, N, O, W	10 słów
A, I, K, L, M, N, O	10 słów

Rysunek 5: Najlepsze kombinacje liter dla Słownika alternatywnego.