

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

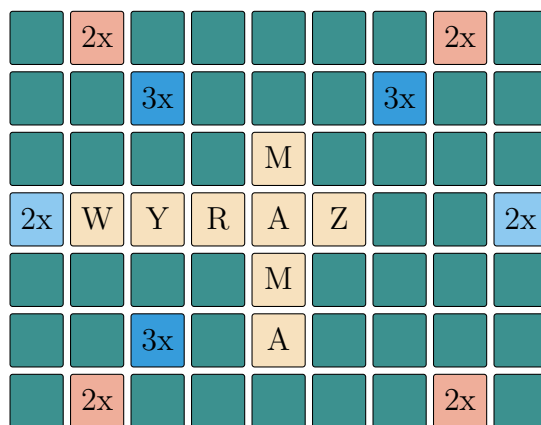
Scrabble to „gra słowna polegająca na układaniu na określonej planszy wyrazów z losowanych liter”¹. Jest to bardzo ogólna definicja, którą należy uściślić. Scrabble jest grą przeznaczoną dla 2-4 osób. Akcesoriami do gry są: kwadratowa plansza o stałym rozmiarze 15 × 15, torebka wypełniona płytkami, na 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”.



Rysunek 2: Fragment planszy. Ruch przez dołożenie liter A, S, Z tworzy dwa wyrazy w układzie krzyżówkowym.

W trakcie jednego ruchu zawodnik może układać płytki tylko w jednym kierunku (pionowo lub poziomo). Utworzony przez zawodnika wyraz musi być spójny, to znaczy, że wszystkie płytki muszą przylegać do siebie bezpośrednio lub poprzez płytki już istniejące na planszy. Wymagane jest, aby tworzone słowo przylegało do przynajmniej jednej płytki, która jest już umieszczona na planszy (nie dotyczy to pierwszego ruchu).

Punktacja za dane zagranie jest obliczana jako suma punktów za wszystkie płytki, które wchodzi w skład utworzonych wyrazów (a więc również tych, które przed zagranie znajdowały się na planszy), z uwzględnieniem niewykorzystanych premii wynikających z pozycji płytki na planszy:

Premia literowa Podwaja lub potraja wartość danej płytki.

Premia słowna Podwaja lub potraja wartość całego wyrazu.

Słowniki do gier

Reguły Scrabble dopuszczają układanie „wszystkich słów występujących w słownikach języka polskiego oraz wszystkich ich

prawidłowych form gramatycznych, z wyjątkiem takich słów, które rozpoczynają się od wielkiej litery, są skrótami, bądź słowami wymagającymi cudzysłowu lub łącznika”².

Brak zamkniętej listy słów możliwych do wykorzystania w grze prowadzi do konfliktu interesów, ponieważ gracze sami muszą rozstrzygnąć między sobą, czy dany wyraz jest legalny, bądź nie. Aby możliwa była uczciwa rozgrywka, konieczne było stworzenie słownika do gier. Słownik wyrazów do gier jest to lista wszystkich słów, wraz ze wszystkimi poprawnymi odmianami, dopuszczalnych do wykorzystania w grach słownych³.

W chwili obecnej istnieją dwa duże polskie słowniki do gier:

Oficjalny Słownik Polskiego Scrabblisty

Przygotowany przez wydawnictwo naukowe PWN. Dopuszcza użycie wyrazów znajdujących się w słownikach PWN wydanych po roku 1980.

Słownik alternatywny Przygotowany

na potrzeby serwisu z grami przez Internet <http://kurnik.pl>. Jest rozwijany przez administratorów serwisu przy współpracy z internautami. Dopuszcza użycie wyrazów znajdujących się w słownikach dowolnego wydawnictwa wydanych po roku 1980, z wyłączeniem czarnej listy słowników⁴.

Poza dopuszczalnymi źródłami informacji, słowniki posiadają jeszcze jedną znaczącą różnicę. Oficjalny Słownik Polskiego Scrabblisty jest dystrybuowany jako program umieszczony na płycie CD, który nie

²Reguły gry w Scrabble z witryny Polskiej Federacji Scrabble - <http://www.pfs.org.pl/zds.php>.

³Anna Andrzejczuk, *Słowniki do gier słownych jako nowy typ wydawnictw leksykograficznych*. Praca magisterska pod kierownictwem dr hab. R. Pa-welec.

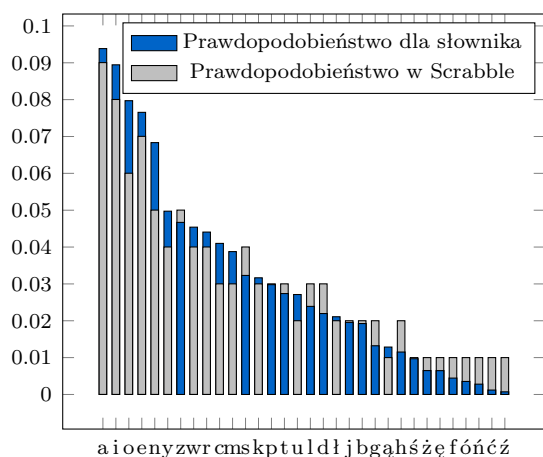
⁴Jest to lista słowników, które nie zostały dopuszczone jako źródło wyrazów ze względu na niedostateczną jakość opracowania.

udostępnia listy wszystkich wyrazów zawartych w słowniku. Program pozwala wyłączyć 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.

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. Korzytnie jest wybierać zagrania, które powodują wyłożenie płytek z literą **h**.



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