Relazione
Esercitazione 6

Frutta con la M

Introduzione

La seguente esercitazione ha come consegna quella di sviluppare un sistema che identifichi un certo termine appartenente ad una certa categoria, data una lettera iniziale.

Durante l'esecuzione viene chiesto all'utente di inserire un argomento (categoria) e una lettera (lettera iniziale), a questo punto il programma svolgerà una ricerca su WordNet per trovare il termine che soddisfa l'input iniziale.

Successivamente viene chiesto all'utente se vuole vedere i successivi risultati (altri termini potenzialmente adeguati), altrimenti il programma termina.

NOTE:

Non c'è competizione con l'utente, per cui il programma è stato pensato per svolgere il ruolo di *aiutante* quando all'utente serve per l'appunto aiuto nell'individuazione di una certa parola.

Struttura del codice ed implementazione

Il codice è presente nel file **`main.py`** che si occupa di richiamare i metodi principali per l'esecuzione del task, ovvero

- `get_topic_synsets(word)`
  metodo per recuperare i synsets di una parola in input
- `get_hyponyms(synset)`
  metodo che, dato un synset in input, ritorna la gerarchia (annidata) di tutti i suoi iponimi.
- `rank_hyponyms(hyponyms_list, synset)`

metodo che data una lista di iponimi (synset) e un synset di riferimento ritorna la distanza (shortest path) tra ogni synset della lista di iponimi e il synset di riferimento. Il valore della distanza è visto come un `rank`: più la distanza è minore più il synset della lista degli iponimi sarà *vicino* al synset di riferimento (e quindi attendibile).
La lista viene poi restituita ordinata dal valore più piccolo del `rank`.

- `synset_to_lemma(synsets_list)`
  metodo che data una lista di synset ritorna una lista dei lemma dei synset (mantenendo il `rank`)

- `find_word_by_start_letter(word_list, letter)`
  controlla se nella lista di parole data in input ci sia una parola che inizi con la lettera data in input

Risultati

Di seguito vengono riportate tre schermate, relative a diversi input

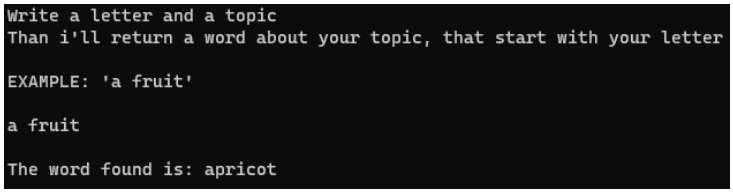

Figura 1 categoria: frutta – lettera A

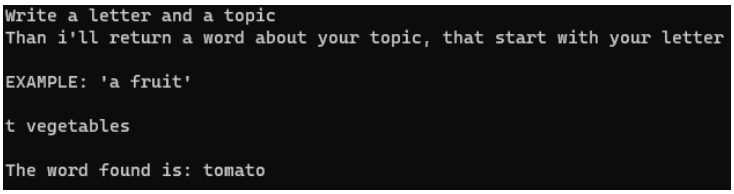

Figura 2 categoria: verdura - lettera: T

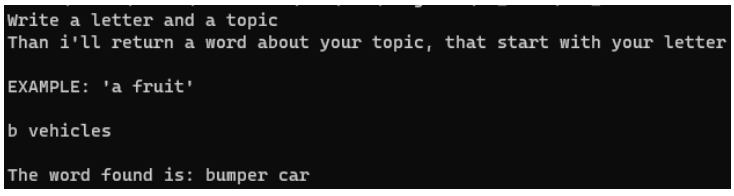

Figura 3 categoria: veicoli - lettera: B

Conclusioni

È necessario indicare una criticità: seppur i risultati ottenuti siano accettabili, per certi argomenti (categorie) questo programma non riesce a fornire sempre un corretto output, questo per via dei synset di WordNet e come sono organizzati.

Infatti, in questo programma si cercano i termini a partire dal primo synset di un argomento e non anche nei successivi.

Es

```
wn.synsets('fruit') -> [Synset('fruit.n.01'), Synset('yield.n.03'), Synset('fruit.n.03'),
Synset('fruit.v.01'),  Synset('fruit.v.02')]
```

Nel caso di *fruit*, il programma funziona bene perché gli iponimi sono presenti solo nel primo synset ovvero `fruit.n.01`, in altri casi invece sono presenti anche nei successivi.

Ciò nonostante, risulta complesso implementare un controllo per i successivi synset perché andrebbe a compromettere il *rank* definito nel programma, non riuscendo a identificare quale valore associare per ogni iponimo preso da ogni lista di iponimi dei synset.