

Reshape Prose into Poetry

Luke Allen
source@stanford.edu

Rhyme scoring

- CMU pronunciation dictionary (included in Python NLTK) lists word pronunciations.

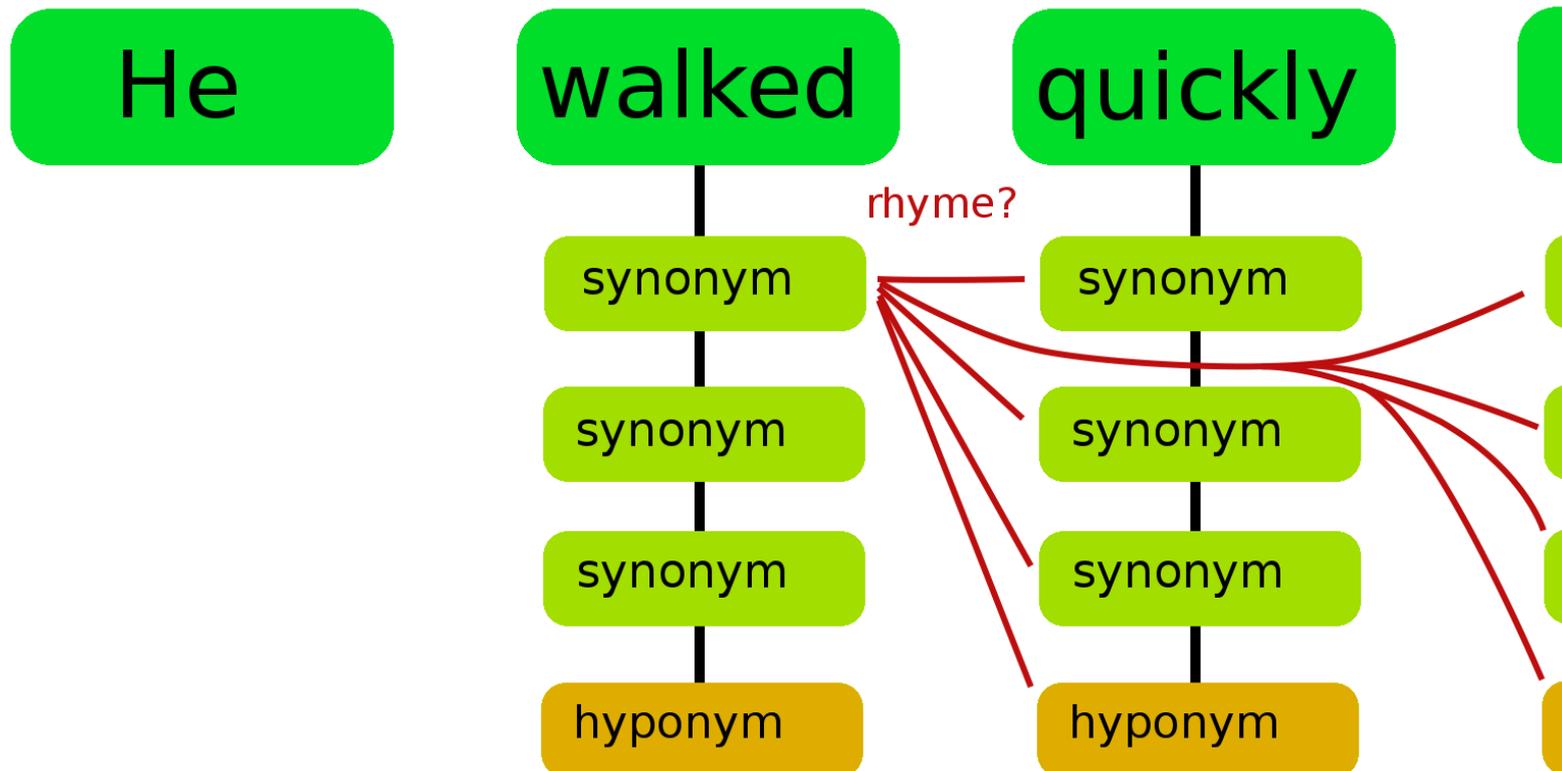
scar => S K AA1 R

bar => B AA1 R

- Compare syllables going backward. More matching syllables = better rhyme. If first consonant and vowel don't match, it's not a rhyme at all.

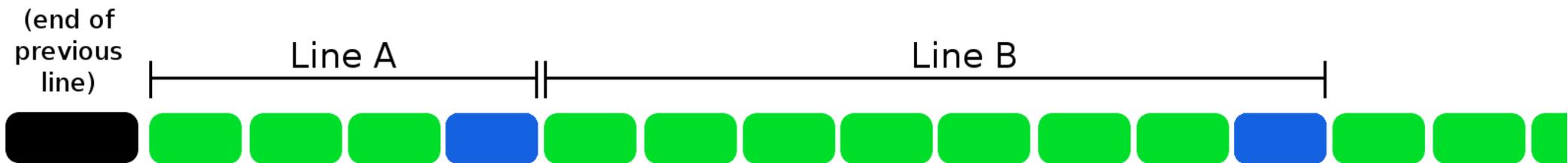
Find Possible Rhymes

- Read in a sample of prose
- Use WordNet synonyms/hyponyms/hypernyms to make a list of acceptable replacements for each word in the original text
- Check each list against every other list (within some distance limit) to find possible rhymes



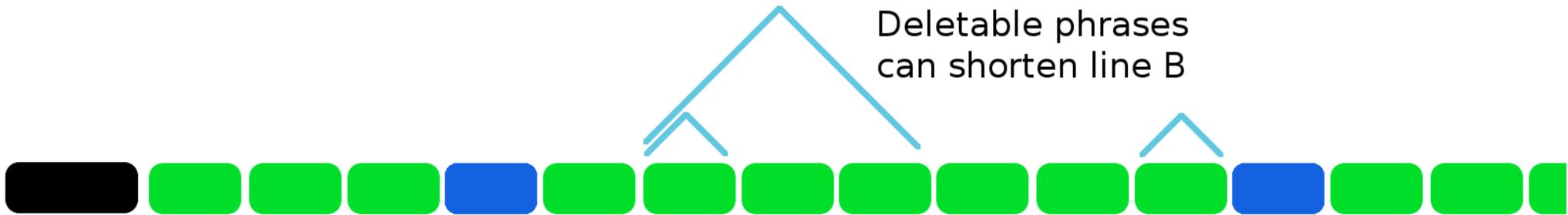
Rhyming lines

- Now we have pairs of rhyming words identified. Each rhyme pair would form two rhyming lines, if we just inserted a newline after each word:



- But the lines are probably different lengths, which is unacceptable for poetry
- Need to equalize lengths

Shortening a line



- Some parts of speech, like adjectives and adverbs, can usually be removed
- When I parsed the original text, I tagged these removable phrases
- Search all combinations of removable phrases, in both lines, to equalize number of syllables in the two lines

Results

--Input--

The ultimate cause of our failure was a simple one: despite all statements to the contrary, it was not due to lack of bravery on the part of our men, or to any fault of the Fleet's. We were defeated by one thing only-by the inferior science of our enemies. I repeat-by the inferior science of our enemies.

When the war opened we had no doubt of our ultimate victory. The combined fleets of our allies greatly exceeded in number and armament those which the enemy could muster against us, and in almost all branches of military science we were their superiors. We were sure that we could maintain this superiority. Our belief proved, alas, to be only too well founded.

(from the short story *Superiority* by Arthur C. Clarke)

--output--

Of our failure the ultimate cause was a simple one : of the Fleet 's, it was not to lack, or to any fault. When the war capsize we had no doubt. The combined fleets of our allies exceeded those which the enemy could muster against us, and of scientific agriculture in all branches we were their leader. Our psychological feature proved alas to be only too well founded ***<--no rhymes

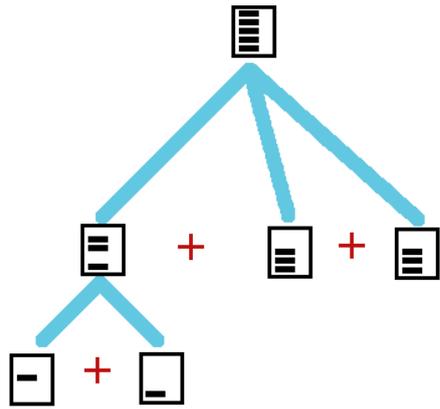
(processing time: 10 seconds)

Limitations:

- Needs WSD for choosing good synonyms
- Can't tell when a hyponym is appropriate
- Some prepositional phrases shouldn't be deleted

Bonus slides

Why checking all possible deletions can be done in polynomial time



- If deletable words are found inside a larger deletable phrase, they are stored as child nodes.
- Each node combines its children with each other and itself, to form a menu of deletion options.
- Fortunately, we can limit the number of combinations to consider: index the combinations by the number of syllables that they delete, and only keep the best combination for each number of syllables. Also discard combinations which delete more syllables than we're interested in.

How to combine deletion options:

Existing deletion options

- 5: [['quite'],['under', 'the','sky']]
- 4: [['under', 'the','sky']]
- 2: [['quite'],['good']]
- 1: [['good']]
- 0: [[]]

+

Independent new options

- 4: [['very'], ['happy']]
- 2: [['very']]
- 0: [[]]

Combine each entry of first list with each entry of second.
Only keep the best option for each number of syllables

- 5: [['quite'],['under', 'the','sky']]
- 4: [['under', 'the','sky']]
- 3: [['good'],['very']]
- 2: [['very']]
- 1: [['good']]
- 0: [[]]

Key idea: only need to combine overall entries; no need to consider breaking apart entries. If the individual components were useful, they would be duplicated in their own entry.