



# Nehovah: A Neologism Creator Nomen Ipsum

Michael R. Smith, Ryan Hintze and Dan Ventura  
Brigham Young University  
Department of Computer Science

# Neologisms

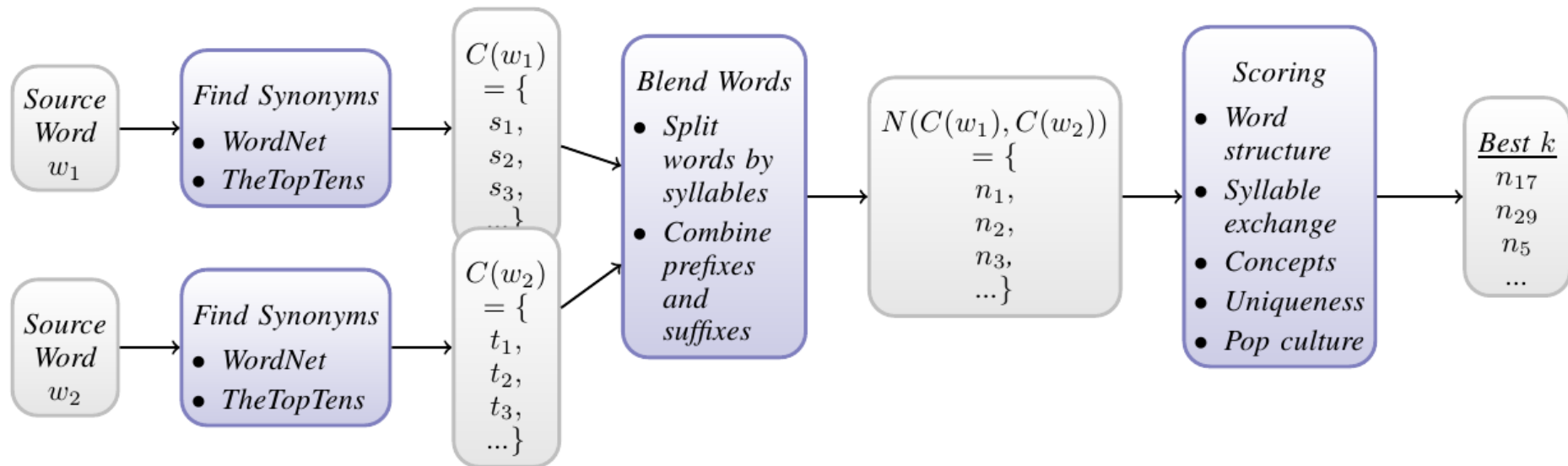
- Language is constantly evolving:
  - Recently added words to the Merriam-Webster Dictionary:
    - Freegan: “an activist who scavenges for free food (as in waste receptacles at stores and restaurants) as a means of reducing consumption of resources”
    - Turducken: “a boneless chicken stuffed into a boneless duck stuffed into a boneless turkey”
    - Gamification: “the process of adding games or gamelike elements to something (as a task) so as to encourage participation”
  - Especially important to many businesses
    - 433,651 trademark applications filed in 2013
    - 4.5% increase from 2012
  - Literary devices
    - Convey meaning – “cybernetics”
    - Add humor and interest as done by Dr. Seuss

# Outline

- Overview of Nehovah
- Synonym and hyponym discovery
- Blending words
- Scoring
  - Word structure scoring
  - Concept scoring
  - Uniqueness scoring
  - Pop culture scoring
- Combining Scores

# Overview of Nehovah

- Goal: Create novel words expressing the concepts from multiple source words



- Synonym and hyponym discovery
- Blending words
- Evaluation

# Synonym and Hyponym Discovery

## ■ Soda

### ■ WordNet

2 senses of soda

Sense 1

sodium carbonate, washing soda, sal soda, soda ash, soda

=> salt

Sense 2

pop, soda, soda pop, soda water, tonic

=> soft drink

## ■ Addiction

### ■ WordNet

3 senses of addiction

Sense 1

addiction, dependence, dependance, dependency, habituation

=> physical condition, physiological state, physiological condition

Sense 2

addiction

=> craving

Sense 3

addiction

=> award, awarding

# Synonym and Hyponym Discovery

## ■ Soda

### ■ TheTopTens

- Top Ten Sodas
  - Coca Cola Cherry
  - Coca Cola
  - Berries and Cream Dr. Pepper
  - ...
- Best Soda Flavors
  - Dr. Pepper
  - Root Beer
  - Mountain Dew
  - ...
- Top 10 Sodas Mexico vs U.S.A.
  - A&W Root Beer (U.S.)
  - Peñafiel (Mexico)

## ■ Addiction

### ■ TheTopTens

- Greatest Janes Addiction Songs
  - Ted, Just Admit It...
  - Been Caught Stealing
  - Stop
  - Mountain Song
  - ...

Validated against Wikipedia

# Blending Words

carbonate			physiological	
Prefix	Suffix		Prefix	Suffix
--	carbonate		--	physiological
ca	rbonate		phy	siological
car	bonate		physi	ological
carbo	nate		physio	logical
carbona	te		physiolo	gical
carbonate	--		physiologi	cal
			physiological	--

caological

carbological

carbonacal

physite

physiolobonate

physiologicalte

# Scoring

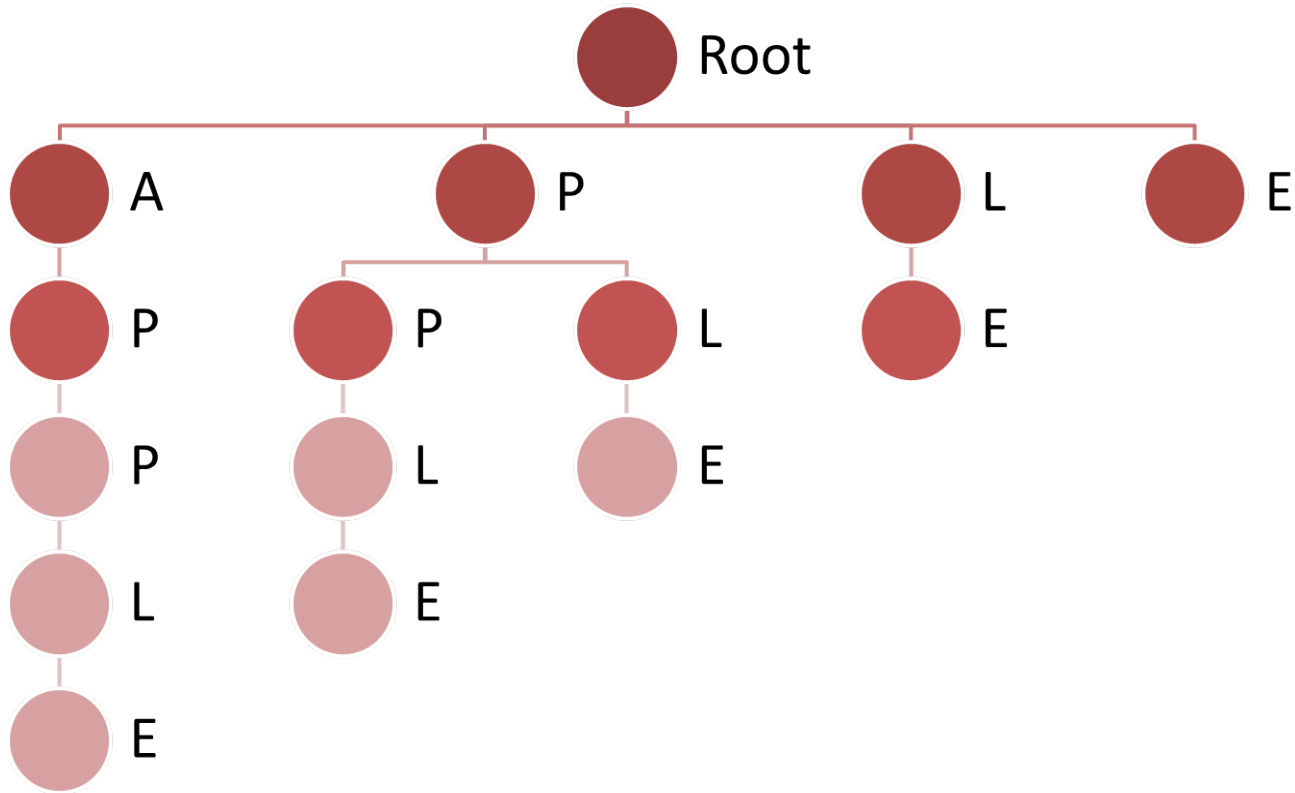
- **Word Structure**
  - Is the neologism realistic?
- **Concept**
  - How apparent are the concepts from the source words?
- **Uniqueness**
  - How unique are the words used to create the neologism?
- **Pop Culture**
  - Are pop culture references used to create the neologism?



# Word Structure Scoring

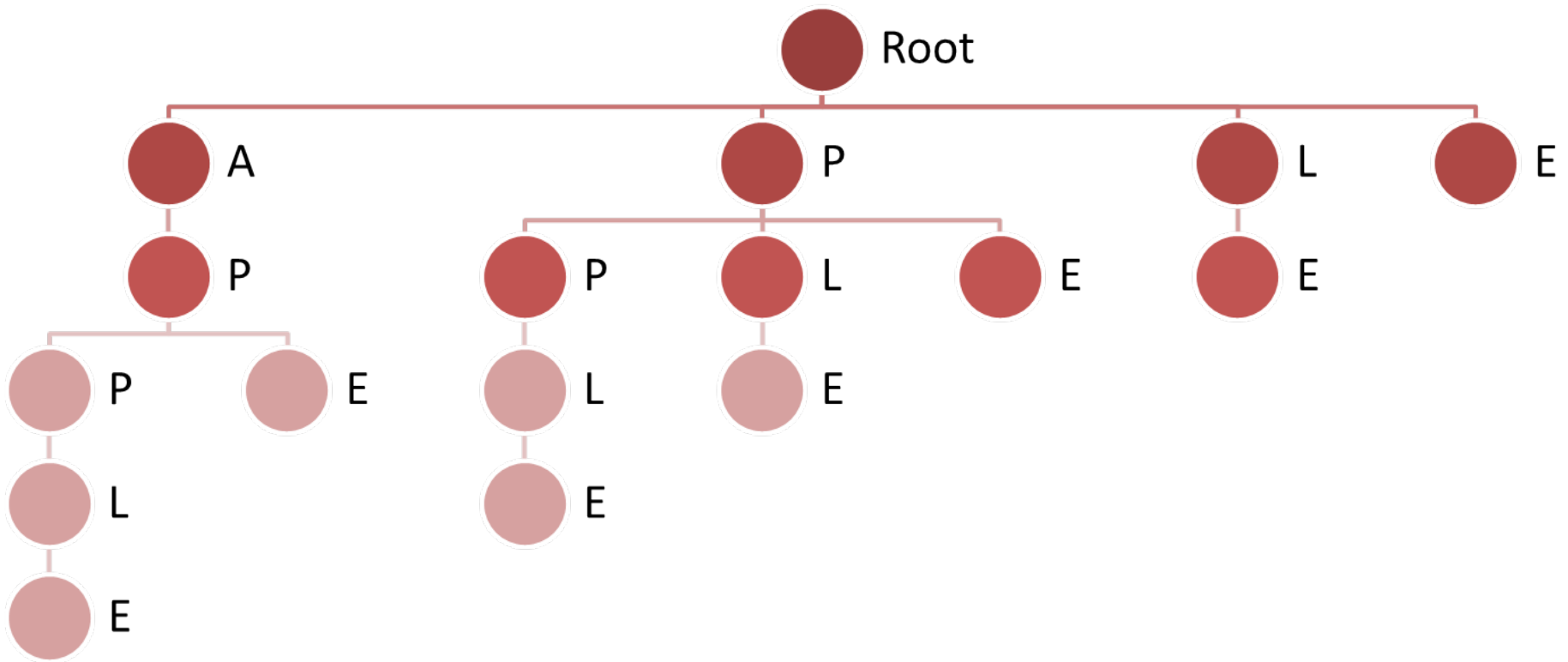
- Valid letter sequence
  - Create an N-gram tree
    - Any Sequence of N letters within a word
    - Useful for confirming that a word is realistic
    - Rejects neologisms with invalid letter sequences
- Syllable exchange/Overlapping letters
  - Does the neologism maintain the flow of one of the source words?
    - Ginormous (from giant and enormous)
    - Linsanity (from Lin and insanity)

# N-Gram Tree



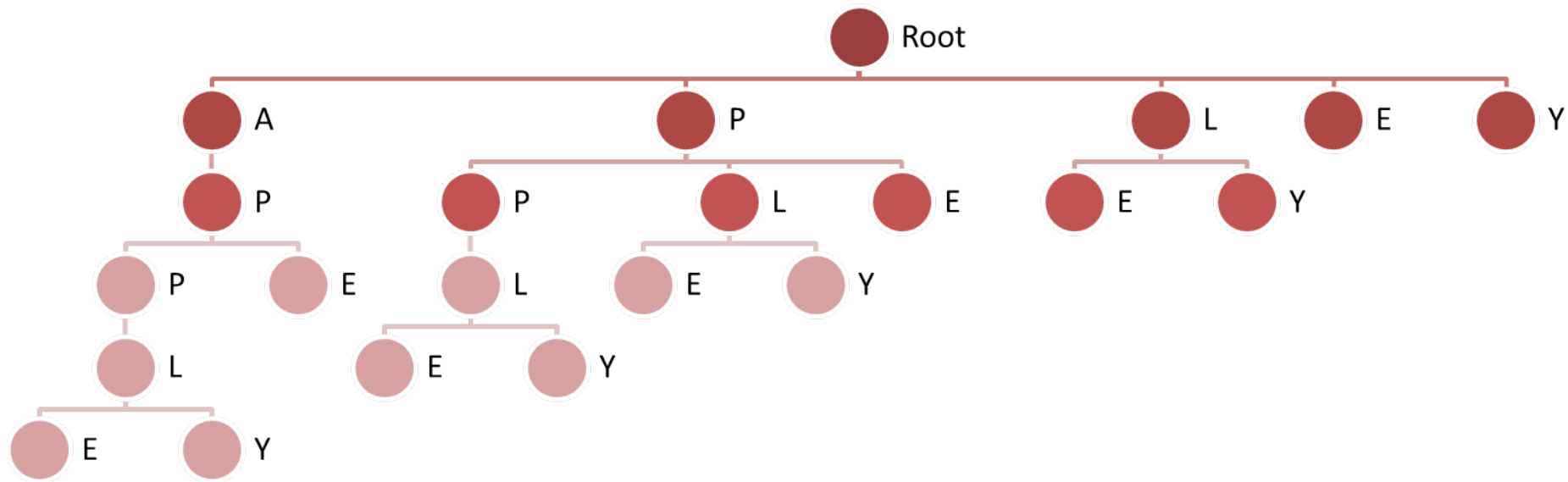
# N-Gram Tree

- Added "Ape"



# N-Gram Tree

- Added "Apply"



If letter sequence is not in the N-gram tree, deemed as invalid

# Word Structure Scoring

- Valid letter sequence
  - Create an N-gram tree
    - Any Sequence of N letters within a word
    - Useful for confirming that a word is realistic
    - Rejects neologisms with invalid letter sequences
- Syllable exchange/Overlapping letters
  - Does the neologism maintain the flow of one of the source words?
    - Ginormous (from giant and enormous)
    - Linsanity (from Lin and insanity)

# Word Structure Scoring

- Valid letter sequence
  - Create an N-gram tree
    - Any Sequence of N letters within a word
    - Useful for confirming that a word is realistic
    - Rejects neologisms with invalid letter sequences
- Syllable exchange/Overlapping letters
  - Does the neologism maintain the flow of one of the source words?
    - Ginormous (from giant and enormous)
    - Linsanity (from Lin and insanity)

# Concept Scoring

## ■ MoreWords

Normalized FPM for words starting/end with a prefix/suffix

### ■ Starting with “carbo”

carbon (50)

carbonate (11)

carbohydrate (8)

carboniferous (4)

carbohydrates (2)

carbonates (1)

### ■ Ending with “ological”

psychological (28)

biological (20)

ideological (17)

technological (16)

geological (10)

...

## ■ Pop Culture

- Normalized count of the number of times a word appears in TheTopTens lists

# Uniqueness Scoring

- For each based work, the relative frequency per million words as provided by MoreWords
  - pop – 79
  - carbonate – 11
  - tonic – 11
  - soda – 8



# Pop Culture Scoring

- Piece-wise equation:
  - 1 if both base words are pop culture words
  - 0.5 if only one is a pop culture word
  - 0 otherwise

# Combining Scores

- A weighted average of the scores
  - Weights are assigned by a user

## NEHOVAH

Word 1:

Word 2:

Pop culture   Concepts   Word Struct   Unique

▶ masterminton  
 Score: 0.9507630344827586

▶ craniatein  
 Score: 0.997815632183908

▶ badminstein  
 Score: 0.9972832183908046

▼ badmintality  
 Score: 0.9649412528735631  
 Alternate Forms: badmintality  
 Prefix Root: badminton  
 Suffix Roots: mentality  
 Pop Culture score: 0.6  
 Concepts Score: 1  
 Word Structure Score: 0.877663

▶ encepberry  
 Score: 0.9986680459770114

▶ badmintein

[BYU - Computer Science](#) - [Machine Learning and Neural Networks Lab](#) - [Suggestions](#)  
 Created by: [Michael R. Smith](#) and Ryan S. Hintze



# Examples

Neologism	Base words		Source Words	
Nehovah	neologism	Jehovah	neologism	creator
divinage	divine	coinage	neologism	creator
machinative	machine	creative	machine	creative
Spritependency	Sprite	dependency	soda	addiction
Pepsidiction	Pepsi	addiction	soda	addiction
pisome	pie	awesome	awesome	pizza
pimazing	pie	amazing	awesome	pizza
iniquitivate	iniquity	cultivate	evil	school
immoralize	immorality	civilize	evil	school
coalesception	coalesce	conception	concept	blend
portmanception	portmanteau	conception	concept	blend

# Future Work

- More diverse source for synonyms/hyponyms
  - Twitter
- Improved word splitting
- Have Nehovah learn which neologisms are best
- Generate a definition for each neologism