TACKLING NATURAL LANGUAGE GENERATION CHALLENGES AT NARRATIVE SCIENCE

Mike Pham & Clayton Norris

Oct 19, 2017

Narrative Science

OVERVIEW

WHAT IS QUILL?

Quill is an Advanced Natural Language Generation (NLG) platform

NLG A form of artificial intelligence (AI) that automatically produces language from structured data.

intent-driven Advanced NLG uses intent, or what you want to know, as its guide from the very beginning.

HOW IS THIS DIFFERENT THAN OTHER NLG?

So what?

HOW IS THIS DIFFERENT THAN OTHER NLG?

So what?

- How is this different than Amazon sending me a templated email receipt of my recent purchases?

HOW IS THIS DIFFERENT THAN OTHER NLG?

So what?

- How is this different than Amazon sending me a templated email receipt of my recent purchases?
- What about all those neural nets generating facebook posts that sound eeriely like my previous posts?

HOW IS THIS DIFFERENT THAN OTHER NLG?

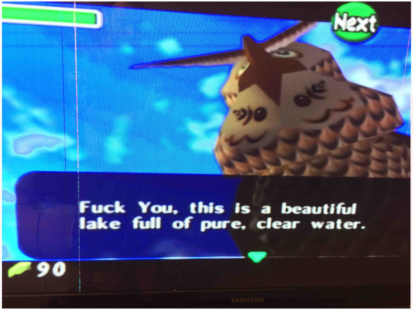
So what?

- How is this different than Amazon sending me a templated email receipt of my recent purchases?
- What about all those neural nets generating facebook posts that sound eeriely like my previous posts?

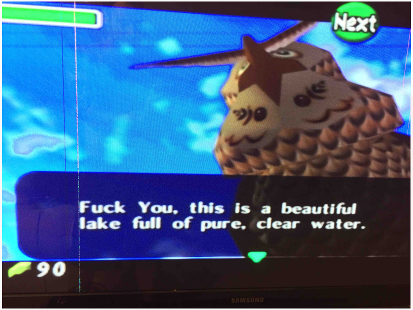
OTHER NLG

trigger warning: offensive language

trigger warning: offensive language



trigger warning: offensive language



- What seems off here?

- Video games conversations have complex decision trees

LIMITATIONS OF NLG

- Video games conversations have complex decision trees
    - Can result in very good and/or appropriate language
    - …but often is mad-libby
    - Flexibility and linguistic creativity is limited and/or unscaleable in production

LIMITATIONS OF NLG

- Video games conversations have complex decision trees
  - Can result in very good and/or appropriate language
  - …but often is mad-libby
  - Flexibility and linguistic creativity is limited and/or unscaleable in production
- Neural nets can learn from data to generate new language

LIMITATIONS OF NLG

- Video games conversations have complex decision trees
    - Can result in very good and/or appropriate language
    - …but often is mad-libby
    - Flexibility and linguistic creativity is limited and/or unscaleable in production
- Neural nets can learn from data to generate new language
    - Can often produce highly natural and nuanced language
    - but has no idea what it’s saying
    - and we have no idea why it’s saying it either

linguistically savvy & intent-driven

3 STRATEGIES

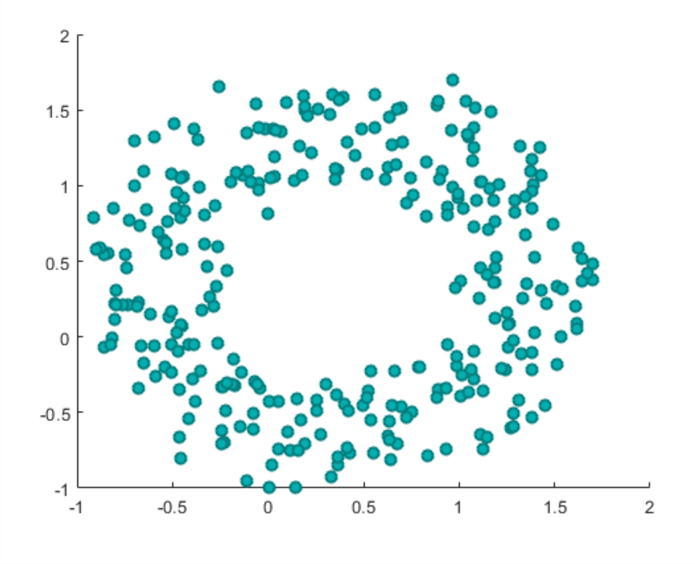
- hardcode/exhaustive listing
- rules/principled based
- machine learning

3 STRATEGIES

- hardcode/exhaustive listing
- rules/principled based
- machine learning

Where does each strategy fit best? How to combine them?

What do you see? How would you recreate this data distribution?



OUTLINE OF TALK

Overview

Irregular Verbs

Pronouns

Sentence Selection

Conclusion

IRREGULAR VERBS

- A single verb can have various **word forms**:

(1) CREATE

a. create, creates, created, creating

b. creator, creation, creative, creatively

- (1a) is an example of **inflectional morphology**
  - expresses grammatical features
  - (usually) doesn't change basic meaning or part of speech

GRAMMATICAL FEATURES

- **Grammatical features** are properties that the grammar of any language tracks and manifests
- Some features that English is sensitive to:
    - **number**: dog, dogs
    - **tense**: create, created
    - **gender**: he, she
    - **person**: we, yall, they
    - **mass/count**: 3 books, *3 bloods
    - **case**: I, me, my, mine

- Word forms can track multiple features at once
- This can be tracked within an **inflectional paradigm**

CREATE

Present

	singular	plural
1	create	create
2	create	create
3	creates	create

Past

	singular	plural
1	created	created
2	created	created
3	created	created

INFLECTIONAL PARADIGMS

- Word forms can track multiple features at once
- This can be tracked within an **inflectional paradigm**

CREATE

Present

	singular	plural
1	create	create
2	create	create
3	creates	create

Past

	singular	plural
1	created	created
2	created	created
3	created	created

- Only 3rd person singular is different – this looks easy!
  - Just add -s to the 3.sg present form and -d to all past forms!

Unfortunately, we all know there are **irregular verbs** in English

BE

Present

	singular	plural
1	am	are
2	are	are
3	is	are

Past

	singular	plural
1	was	were
2	were	were
3	was	were

IRREGULARS

Unfortunately, we all know there are **irregular verbs** in English

BE

Present

	singular	plural
1	am	are
2	are	are
3	is	are

Past

	singular	plural
1	was	were
2	were	were
3	was	were

- Darn, how do we get am or was from be?

BEST STRATEGY

- There are rules for regular morphology
- Which verbs are irregular seems arbitrary
- How irregular verbs inflect also seems arbitrary
- Rules might be tough to derive
- Machine Learning may work, but do we actually want to overfit data?

- Wikipedia lists about 200 English irregular verbs, including shrive, stave, gild
- This is finite set, and most words aren't even that relevant
- Verb dictionaries exist
- There are subgroups within the irregulars
- It is feasible to exhaustively hardcode a list of all irregulars without rules or ML
- We can exactly fit the data without over- or undergeneralizing

PRONOUNS

Anaphora Expressions that depend on a contextual anteceent for their interpretation

Pronoun A type of anaphor that can replace a **Noun Phrase (NP)** (or Determiner Phrase)

Nominative

	singular	plural
1	I	we
2	you	you/yall/yinz
3	she/he/it	they

Accusative

	singular	plural
1	me	us
2	you	you/yall/yinz
3	her/him/it	them

In later years, holding forth to an interviewer or to an audience of aging fans at a comic book convention, Sam Clay liked to declare, apropos of **his** and Joe Kavalier's greatest creation, that back when **he** was a boy, sealed and hog-tied inside the airtight vessel known as Brooklyn, New York, **he** had been haunted by dreams of Harry Houdini. "To **me**, Clark Kent in a phone booth and Houdini in a packing crate, **they** were one and the same thing,"[...]

-Michael Chabon, The Amazing Adventures of Kavalier & Clay

In later years, holding forth to an interviewer or to an audience of aging fans at a comic book convention, Sam Clay liked to declare, apropos of **his** and Joe Kavalier's greatest creation, that back when **he** was a boy, sealed and hog-tied inside the airtight vessel known as Brooklyn, New York, **he** had been haunted by dreams of Harry Houdini. "To **me**, Clark Kent in a phone booth and Houdini in a packing crate, **they** were one and the same thing,"[...]

-Michael Chabon, The Amazing Adventures of Kavalier & Clay

- Connecting reference between expressions is non-trivial!

- Strawman argument: hardcoding every possible instance to use a pronoun is out

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning?

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning? Is likely possible...

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning? Is likely possible...
  - what are the features we want to track?
  - how arbitrary is the data?

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning? Is likely possible...
    - what are the features we want to track?
    - how arbitrary is the data?
- ”Though this be madness, yet there is method in ’t.”

A PRINCIPLED APPROACH

- The distribution of pronouns is not arbitrary
- We actually probably have a pretty good idea of when we can use pronouns

A PRINCIPLED APPROACH

- The distribution of pronouns is not arbitrary
- We actually probably have a pretty good idea of when we can use pronouns
- They seem to **corefer** with recently mentioned entities of that match their description

- The distribution of pronouns is not arbitrary
- We actually probably have a pretty good idea of when we can use pronouns
- They seem to **corefer** with recently mentioned entities of that match their description
- Let's try a rule:

(2) **Pronoun Rule 1**: If the entity is the same as the most recent entity with the same **features** (person, gender, number), a pronoun can be used

(3) a. Harry was in Gryffindor.
b. **He** was friends with Ron.
c. **He** had a pet rat.

- Who does He in (3c) refer to? Harry or Ron?

DOES IT WORK?

(3) a. Harry was in Gryffindor.

b. **He** was friends with Ron.

c. **He** had a pet rat.

- Who does He in (3c) refer to? Harry or Ron?
- Ron is the most salient singular masculine entity
- so why doesn't He corefer?

DOES IT WORK?

(3) a. Harry was in Gryffindor.
    b. **He** was friends with Ron.
    c. **He** had a pet rat.

- Who does He in (3c) refer to? Harry or Ron?
- Ron is the most salient singular masculine entity
- so why doesn't He corefer?
- It seems like linear order is too simplistic of an approach

- Pronouns (across sentences) are tracking **saliency**

Salient: assumed to be in the **addressee**'s consciousness at the **utterance time**

- Pronouns (across sentences) are tracking **saliency**

Salient: assumed to be in the **addressee**'s consciousness at the **utterance time**

(4) Harry studies at Hogwarts with Ron.

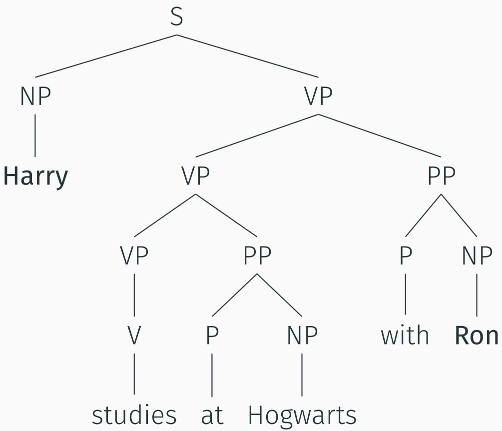
- Who is more salient? Harry? or Ron?

- Pronouns (across sentences) are tracking **saliency**

Salient: assumed to be in the **addressee**'s consciousness at the **utterance time**

(4) Harry studies at Hogwarts with Ron.

- Who is more salient? Harry? or Ron?
- Why?



- Subjects are structurally higher than objects
- In English this correlates with saliency

- An advanced NLG system should track saliency in order to use pronouns
- We've seen that the syntactic structure strongly influences saliency

TRACKING SALIENCY

- An advanced NLG system should track saliency in order to use pronouns
- We've seen that the syntactic structure strongly influences saliency
- Pronoun distribution appears to be based on known principles
- so lwe should ensure that the AI system also shares those principles

- But are there other factors at play?

OTHER FACTORS?

- But are there other factors at play?
    - recency
    - repitition
    - ??

OTHER FACTORS?

- But are there other factors at play?
  - recency
  - repitition
  - ??
- How would these interact with each other?

SENTENCE SELECTION

Sentece generation: only grammatical and accurate sentences should be **generated**

Sentence selection: the stylistically best sentence from the set of grammatical candidate sentences should be **selected**

Sentece generation: only grammatical and accurate sentences should be **generated**

Sentence selection: the stylistically best sentence from the set of grammatical candidate sentences should be **selected**

- but what determines a stylistically ‘good‘ sentence?

WHAT MAKES A GOOD SENTENCE?

- Most native speakers will agree when a sentence is grammatical
- But style is vague and elusive, varying from person to person

- Most native speakers will agree when a sentence is grammatical
- But style is vague and elusive, varying from person to person
- Which do think is the best sentence?

(5)
a. Aaron Young generated $3M in revenue in 2016.
b. Aaron Young's revenue was $3M in 2016.
c. Revenue for Aaron Young was $3M in 2016.
d. In 2016, Aaron Young generated $3M in revenue.
e. Aaron Young's 2016 generated revenue was $3M.

- Most native speakers will agree when a sentence is grammatical
- But style is vague and elusive, varying from person to person
- Which do think is the best sentence?

(5)
a. Aaron Young generated $3M in revenue in 2016.
b. Aaron Young's revenue was $3M in 2016.
c. Revenue for Aaron Young was $3M in 2016.
d. In 2016, Aaron Young generated $3M in revenue.
e. Aaron Young's 2016 generated revenue was $3M.

- is there even a right answer?

- There seem to be multiple factors involved:
    - length
    - subject choice
    - values before attributes
    - fronted information
    - strong verbs vs copulas
    - ...

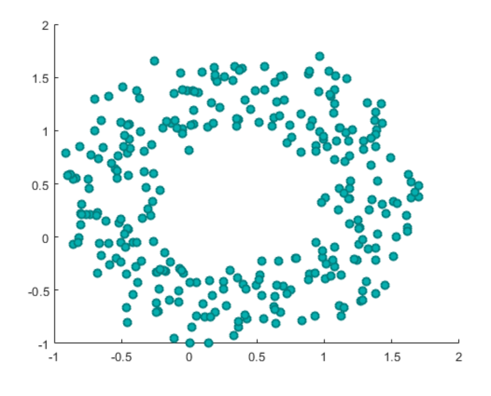
MULTIPLE AXES OF 'GOODNESS'

- There seem to be multiple factors involved:
    - length
    - subject choice
    - values before attributes
    - fronted information
    - strong verbs vs copulas
    - …
- These axes seem largely independent
- Different users also vary in how strongly they weight each factor

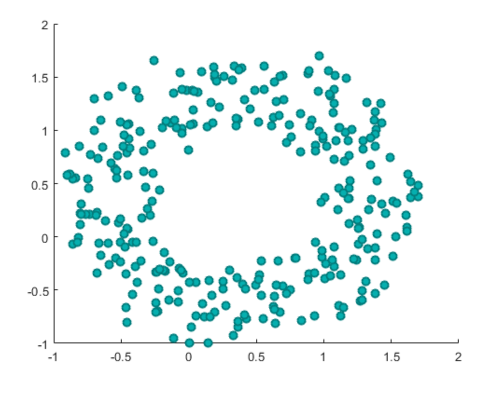
- Sentence selection involves the interaction between several features
- The importance of these features is variable

CONCLUSION

What do you see?

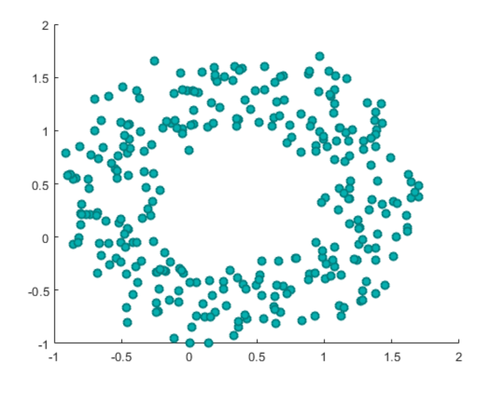


What do you see?



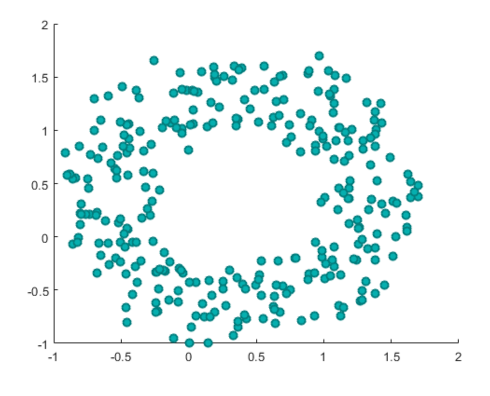
- **Irregular verbs:** discrete points

What do you see?



- **Irregular verbs:** discrete points
- **Pronouns**: conceptual circle → messy data

What do you see?



- **Irregular verbs:** discrete points
- **Pronouns**: conceptual circle → messy data
- **Sentence selection**: messy data → conceptual circle

Get the source of this theme and the demo presentation from

`github.com/matze/mtheme`

The theme itself is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Questions?