

Sheaves and garden-path sentences

Daphne Wang and Mehrnoosh Sadrzadeh
University College London

Pros & Comps
Ljubljana, 2nd August 2023

Sheaves & Garden-path sentences

Sheaves & Garden-path sentences



Sheaves & Garden-path sentences



- ▶ Mathematical object
encoding data **consistent**
with a base structure

Sheaves & Garden-path sentences



- ▶ Mathematical object
encoding data **consistent**
with a base structure
- ▶ Used in topology, logic,
data analysis, quantum
mechanics, ...

Sheaves & Garden-path sentences



- ▶ Mathematical object encoding data **consistent** with a base structure
- ▶ Used in topology, logic, data analysis, quantum mechanics, ...

Sheaves & Garden-path sentences



- ▶ Mathematical object encoding data **consistent** with a base structure
- ▶ Used in topology, logic, data analysis, quantum mechanics, ...
- ▶ Sentences which are hard to parse

Sheaves & Garden-path sentences



- ▶ Mathematical object encoding data **consistent** with a base structure
- ▶ Used in topology, logic, data analysis, quantum mechanics, ...



- ▶ Sentences which are hard to parse
- ▶ Used in psycholinguistics to uncover mechanisms in human parsing



Garden-path sentences

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*
University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

The employees understood the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*
University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

The employees understood the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*
University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

The employees understood the contract would change

Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*

University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

The employees understood the contract would change

Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*

University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

NP/S The employees understood the contract would change

Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*

University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

NP/S The employees understood the contract would change

Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*
University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

NP/**S** The employees understood the contract would change

Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*

University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

NP/S The employees understood the contract would change

NP/Z Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*

University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

NP/S The employees understood the contract would change

NP/Z Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*

University of Connecticut, 1979.

Garden-path sentences

Definition

A garden-path sentence is a sentence which is **grammatically correct** and **unambiguous**, but which forces the reader to initially adopt a **wrong syntactic parse**.

Examples

NP/S The employees understood the contract would change

NP/**Z** Because the employees negotiated the contract would change

Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies.*

University of Connecticut, 1979.

The psycholinguistics theories

The _

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

The employees _



Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories



The employees understood

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

The employees understood the _



Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

The employees understood the contract

The employees understood the contract _

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

very likely



easy to access

The employees understood the contract

The employees understood the contract _

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

very likely



easy to access

The employees understood the contract

unlikely



hard to access

The employees understood the contract _

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

very likely
↓
easy to access

The employees understood the contract ?
would

unlikely
↓
hard to access

The employees understood the contract -

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

impossible

The employees understood the contract ?
would
(cannot be extended)

unlikely

↓
hard to access

The employees understood the contract _

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

impossible

The employees understood the contract ?
would
(cannot be extended)

unlikely

↓
hard to access

The employees understood the contract would ?

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

impossible

The employees understood the contract ?
would
(cannot be extended)

unlikely

↓
hard to access

The employees understood the contract would -
(can be extended)

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

impossible

The employees understood the contract ?
would
(cannot be extended)

has to be adopted

The employees understood the contract would _
(can be extended)

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

The psycholinguistics theories

impossible

The employees understood the contract ?
would
(cannot be extended)

**has to be
adopted**

§

Difficulty

The employees understood the contract would _
(can be extended)

Van Gompel, Roger PG, et al. "Evidence against competition during syntactic ambiguity resolution." *Journal of Memory and Language* 52.2 (2005): 284-307.

Surprisal predictions

Surprisal and reading times

Surprisal predictions

Surprisal and reading times

- ▶ *Observation:* predictable words (in given a context) are read faster than unpredictable ones

Surprisal predictions

Surprisal and reading times

- ▶ *Observation:* predictable words (in given a context) are read faster than unpredictable ones

Definition

The surprisal of a word w_n in the context $w_1 \dots w_{n-1}$ is given by:

$$S(w_n) = -\log_2 (P [w_n | w_1 \dots w_{n-1}]) \quad (1)$$

Surprisal predictions

Surprisal and reading times

- ▶ *Observation:* predictable words (in given a context) are read faster than unpredictable ones

Definition

The surprisal of a word w_n in the context $w_1 \dots w_{n-1}$ is given by:

$$S(w_n) = -\log_2 (P [w_n | w_1 \dots w_{n-1}]) \quad (1)$$

- ▶ It has been shown that reading time (first pass gaze duration) is proportional to surprisal in naturalistic sources (e.g. newspapers)¹

¹ Smith, Nathaniel J., and Roger Levy. "The effect of word predictability on reading time is logarithmic." *Cognition* 128.3 (2013): 302-319.

Surprisal predictions

Surprisal and reading times

- ▶ *Observation:* predictable words (in given a context) are read faster than unpredictable ones

Definition

The surprisal of a word w_n in the context $w_1 \dots w_{n-1}$ is given by:

$$S(w_n) = -\log_2 (P [w_n | w_1 \dots w_{n-1}]) \quad (1)$$

- ▶ It has been shown that reading time (first pass gaze duration) is proportional to surprisal in **naturalistic sources** (e.g. newspapers)¹

¹ Smith, Nathaniel J., and Roger Levy. "The effect of word predictability on reading time is logarithmic." *Cognition* 128.3 (2013): 302-319.

Surprisal predictions

Surprisal in garden-path sentences

Surprisal predictions

Surprisal in garden-path sentences

- ▶ It was first proposed by Hale that surprisal could explain garden-path effects¹

¹ Hale, John. "A probabilistic Earley parser as a psycholinguistic model." *Second meeting of the north american chapter of the association for computational linguistics*. 2001.

Surprisal predictions

Surprisal in garden-path sentences

- ▶ It was first proposed by Hale that surprisal could explain garden-path effects¹
- ▶ Various work² showed that surprisal **underestimates the magnitude** of the garden-path effect and **cannot distinguish** different types of sentences.

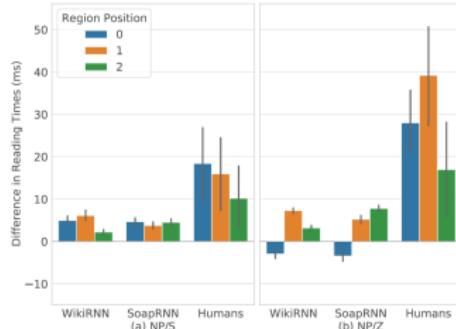
¹ Hale, John. "A probabilistic Earley parser as a psycholinguistic model." *Second meeting of the north american chapter of the association for computational linguistics*. 2001.

²Huang, Kuan-Jung, et al. "Surprisal does not explain syntactic disambiguation difficulty: evidence from a large-scale benchmark." (2023).

Surprisal predictions

Surprisal in garden-path sentences

- ▶ It was first proposed by Hale that surprisal could explain garden-path effects¹
- ▶ Various work² showed that surprisal **underestimates the magnitude of the garden-path effect and cannot distinguish different types of sentences.**



¹ Hale, John. "A probabilistic Earley parser as a psycholinguistic model." *Second meeting of the north american chapter of the association for computational linguistics*. 2001.

²Huang, Kuan-Jung, et al. "Surprisal does not explain syntactic disambiguation difficulty: evidence from a large-scale benchmark." (2023).

Surprisal predictions

Proposed reasons for the failure of surprisal

Surprisal predictions

Proposed reasons for the failure of surprisal

1. Surprisal is not an accurate description of the human processes when garden-path is encountered¹

¹ Van Schijndel, Marten, and Tal Linzen. "Single-stage prediction models do not explain the magnitude of syntactic disambiguation difficulty." *Cognitive science* 45.6 (2021): e12988.

Surprisal predictions

Proposed reasons for the failure of surprisal

1. Surprisal is not an accurate description of the human processes when garden-path is encountered¹
2. Reanalysis is necessary for garden-path resolution (i.e. parsing is not strictly incremental)¹

¹ Van Schijndel, Marten, and Tal Linzen. "Single-stage prediction models do not explain the magnitude of syntactic disambiguation difficulty." *Cognitive science* 45.6 (2021): e12988.

Proposed reasons for the failure of surprisal

1. Surprisal is not an accurate description of the human processes when garden-path is encountered¹
2. Reanalysis is necessary for garden-path resolution (i.e. parsing is not strictly incremental)¹
3. Probabilities extracted from Large Language Models (LLMs) does not accurately represent the probabilities assigned by humans²

¹ Van Schijndel, Marten, and Tal Linzen. "Single-stage prediction models do not explain the magnitude of syntactic disambiguation difficulty." *Cognitive science* 45.6 (2021): e12988.

² Arehalli, Suhas, Brian Dillon, and Tal Linzen. "Syntactic surprisal from neural models predicts, but underestimates, human processing difficulty from syntactic ambiguities." *arXiv preprint arXiv:2210.12187* (2022).



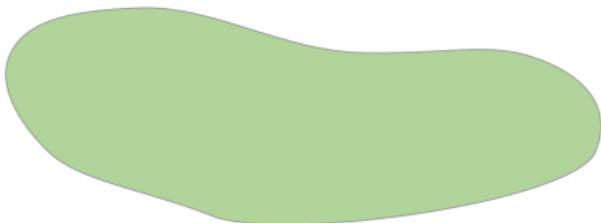
Presheaf models

A crash course in sheaf theory

Preliminaries: Spaces, Restrictions and Presheaves

Context

\mathcal{C}^{op}

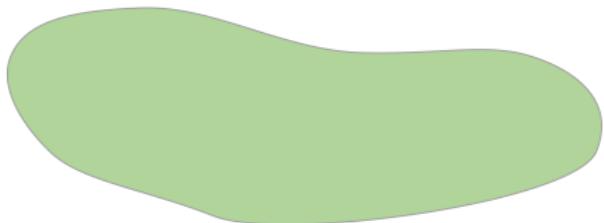


A crash course in sheaf theory

Preliminaries: Spaces, Restrictions and Presheaves

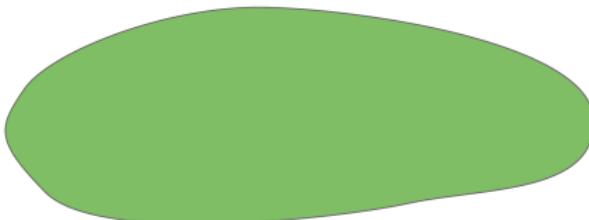
Context

$$\mathcal{C}^{\text{op}}$$



Set

Data



A crash course in sheaf theory

Preliminaries: Spaces, Restrictions and Presheaves

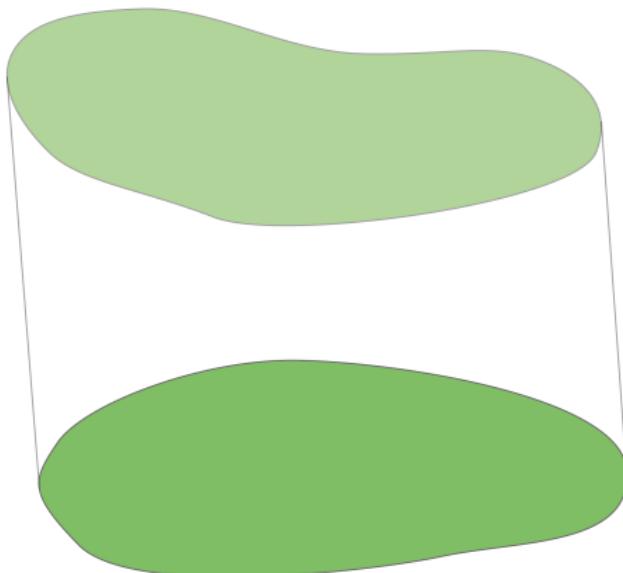
Context

\mathcal{C}^{op}



Set

Data



A crash course in sheaf theory

Preliminaries: Spaces, Restrictions and Presheaves

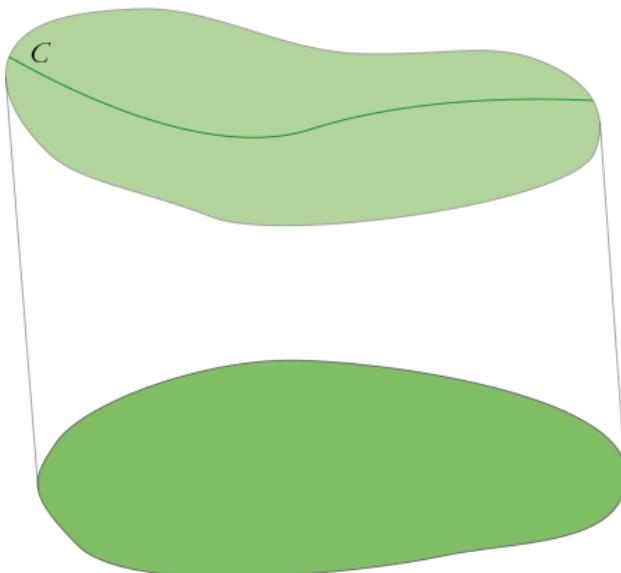
Context

\mathcal{C}^{op}



Set

Data



A crash course in sheaf theory

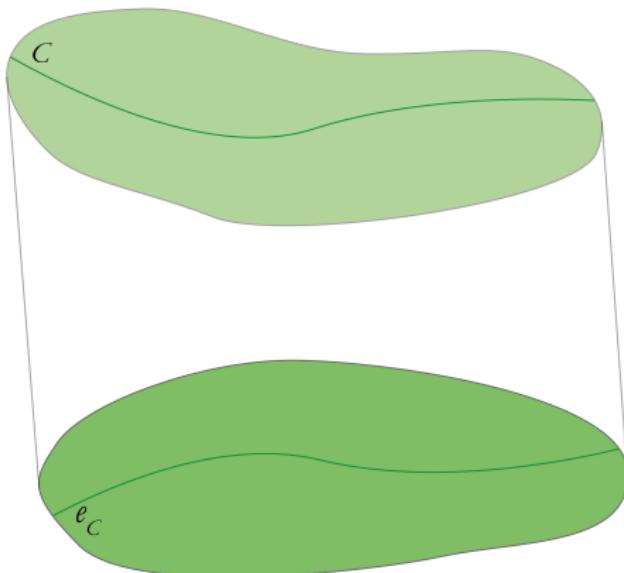
Preliminaries: Spaces, Restrictions and Presheaves

Context

 \mathcal{C}^{op} 

Set

Data



A crash course in sheaf theory

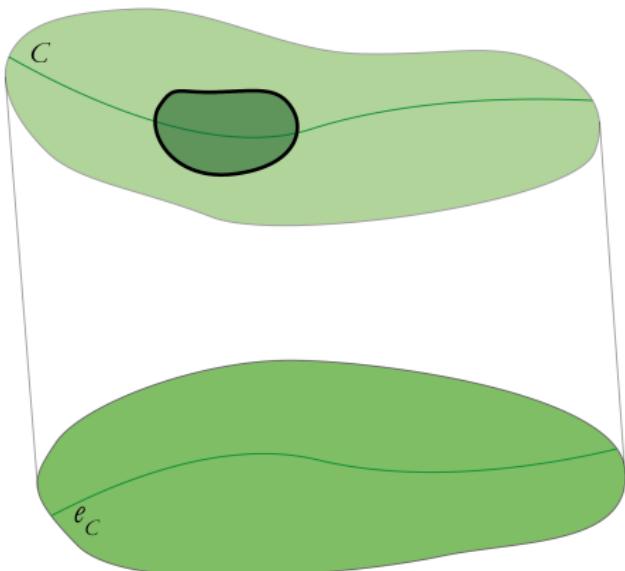
Preliminaries: Spaces, Restrictions and Presheaves

Context

 \mathcal{C}^{op} 

Set

Data



A crash course in sheaf theory

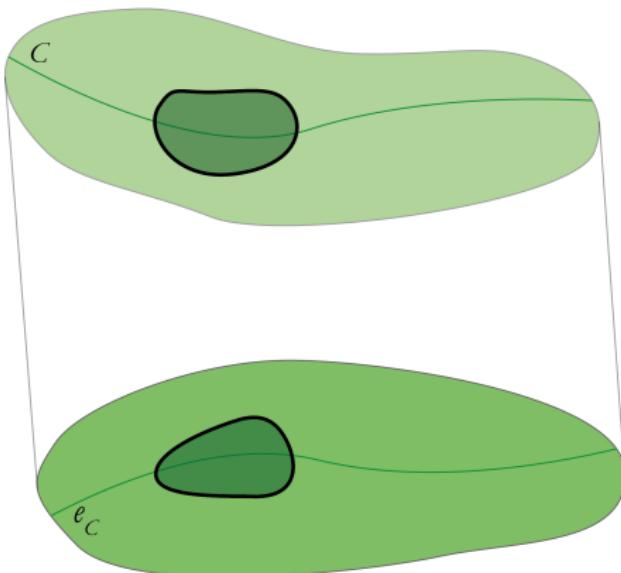
Preliminaries: Spaces, Restrictions and Presheaves

Context

 \mathcal{C}^{op} 

Set

Data



A crash course in sheaf theory

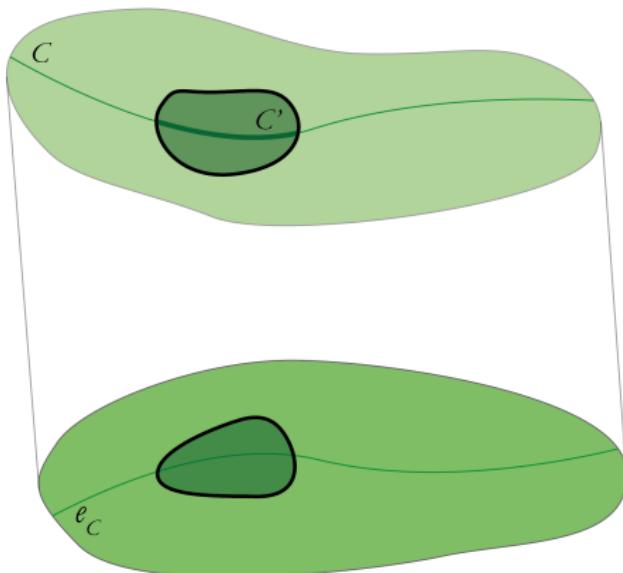
Preliminaries: Spaces, Restrictions and Presheaves

Context

 \mathcal{C}^{op} 

Set

Data



A crash course in sheaf theory

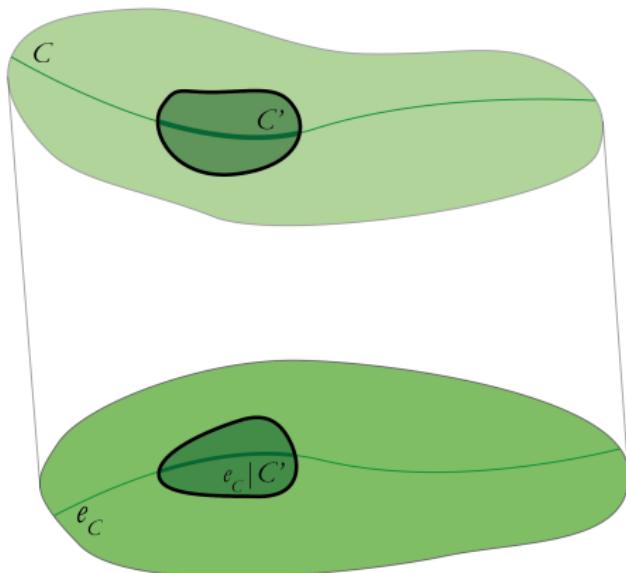
Preliminaries: Spaces, Restrictions and Presheaves

Context

 \mathcal{C}^{op} 

Set

Data



A crash course in sheaf theory

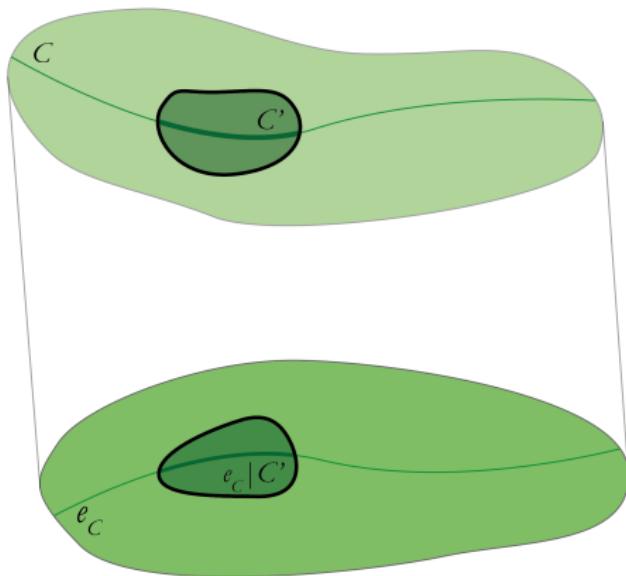
Preliminaries: Spaces, Restrictions and Presheaves

Context

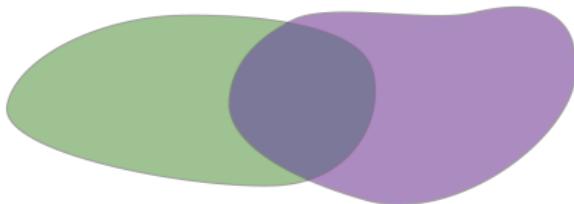
 \mathcal{C}^{op} 

Set

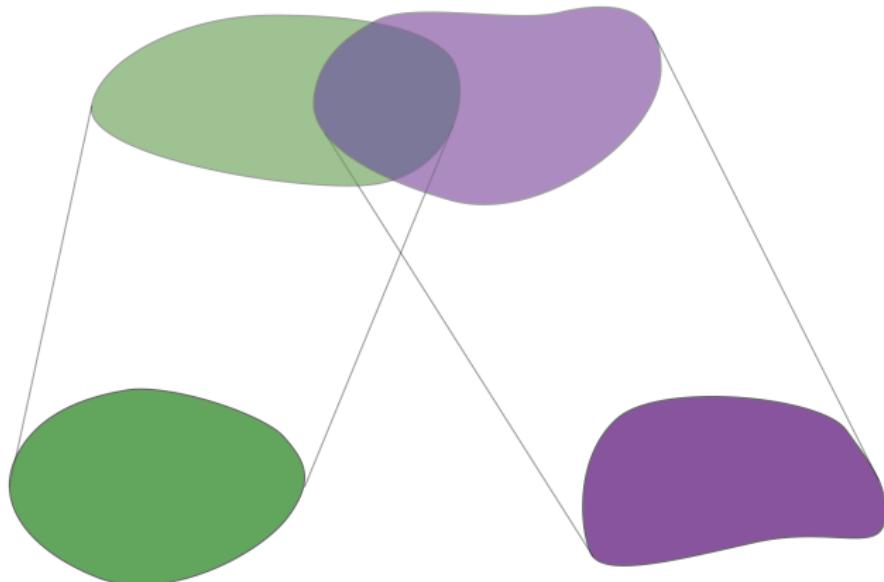
Data

*(Presheaf)*

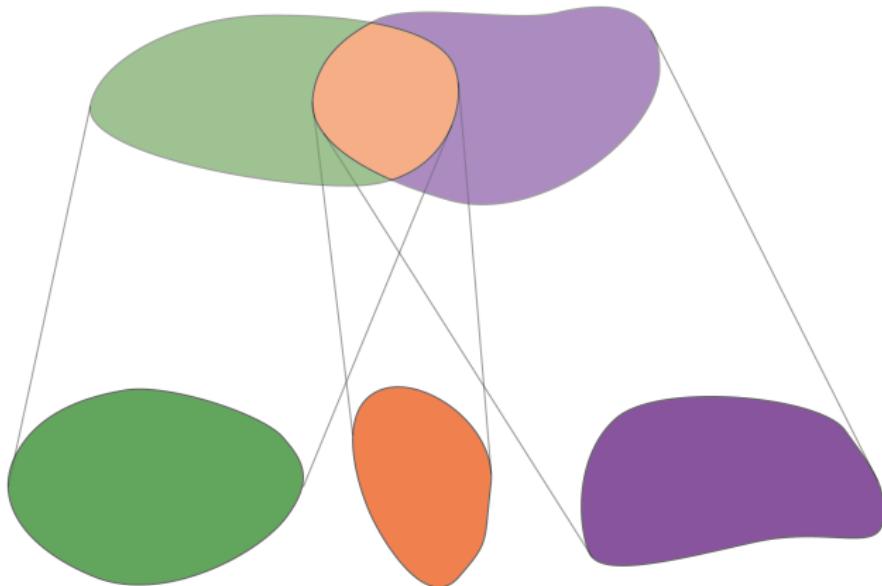
Sheaves



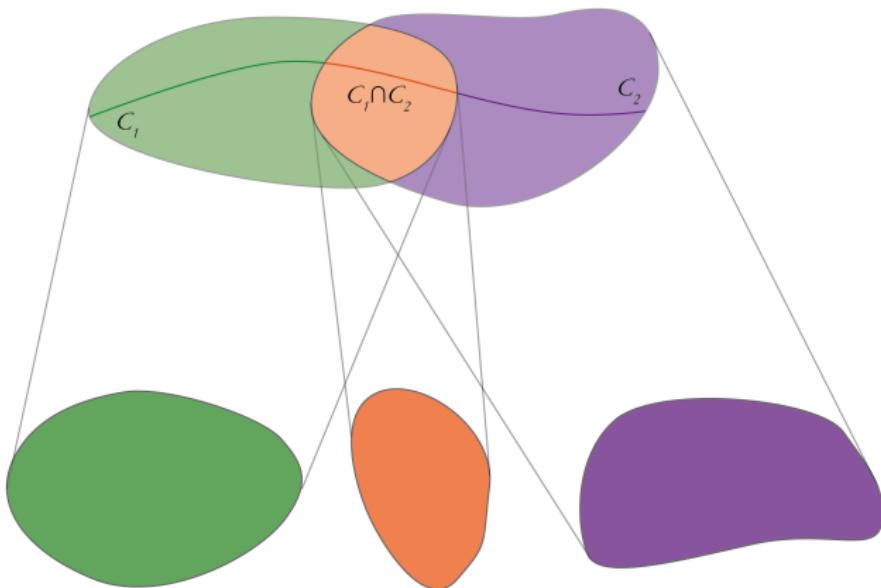
Sheaves



Sheaves

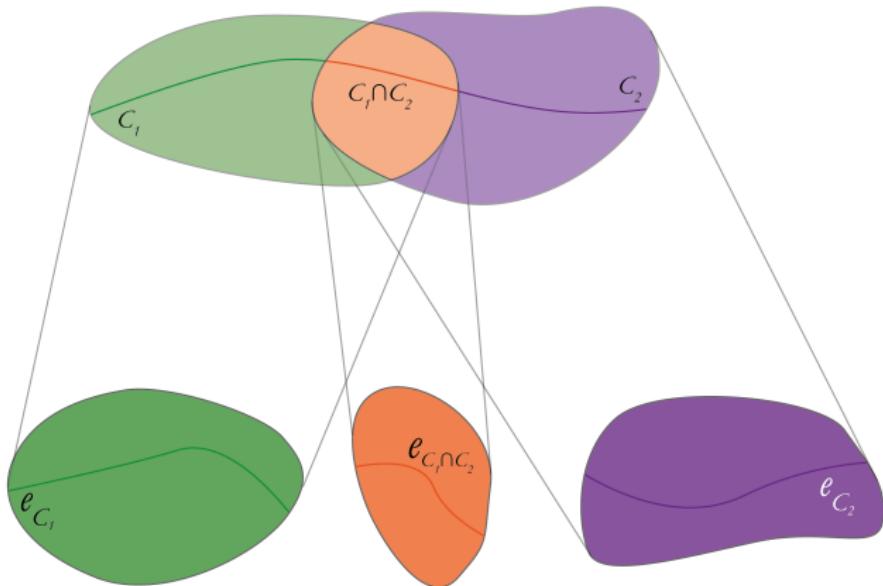


Sheaves

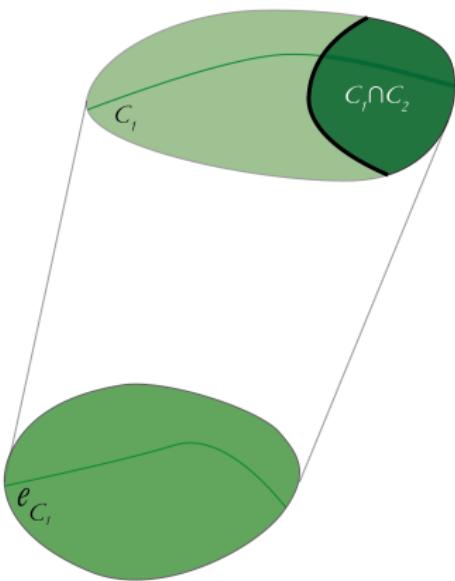


A crash course in sheaf theory

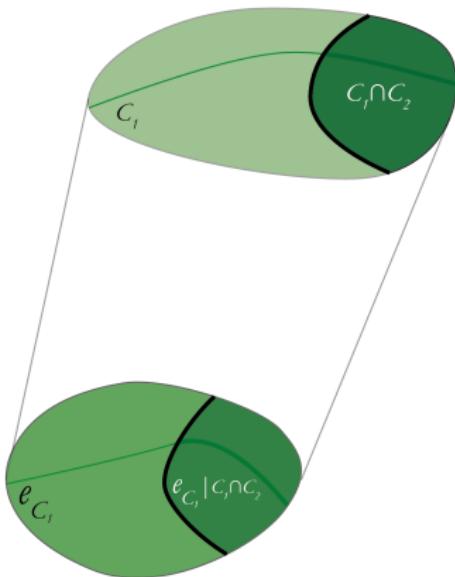
Sheaves



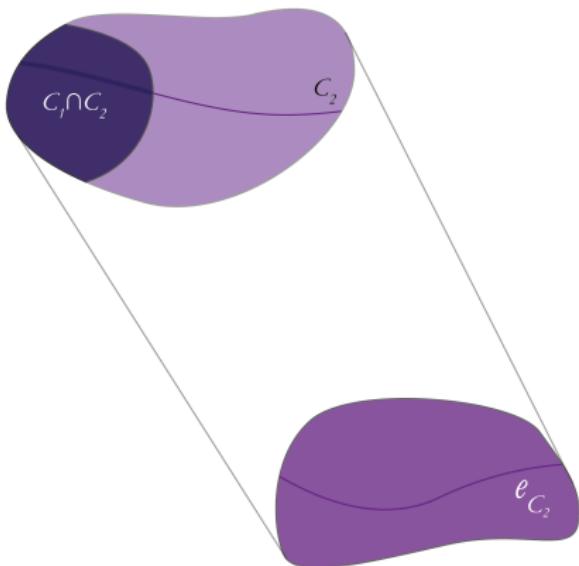
Sheaves



Sheaves

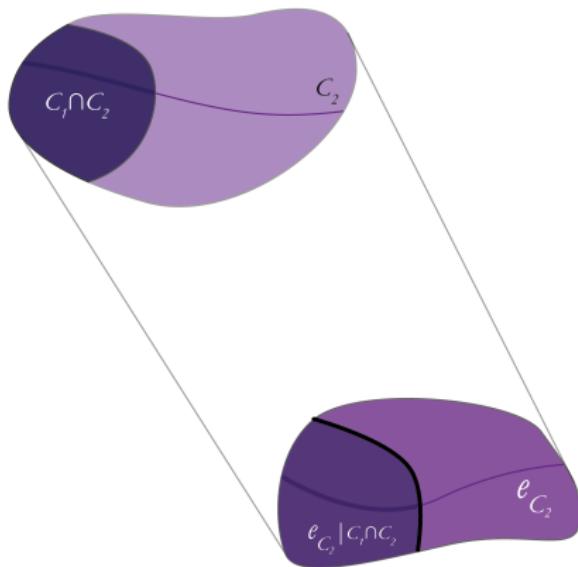


Sheaves



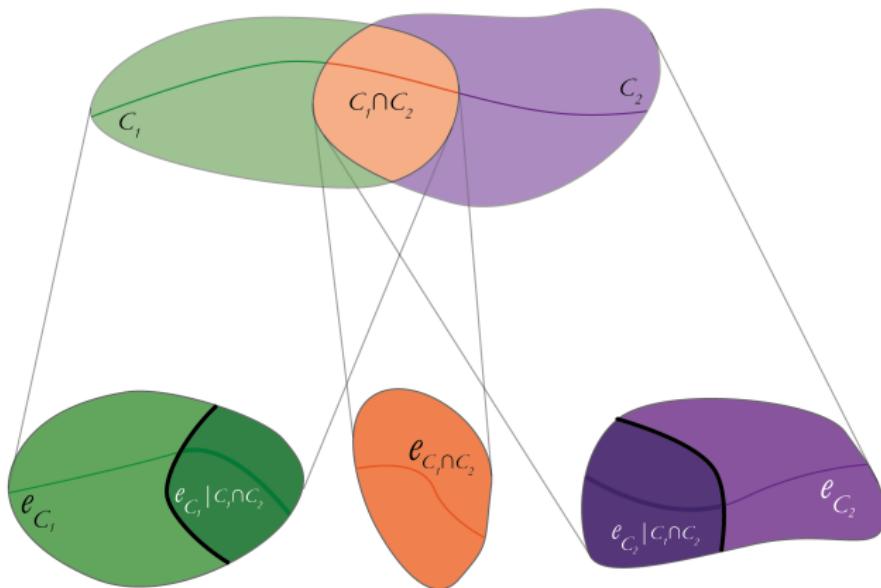
A crash course in sheaf theory

Sheaves



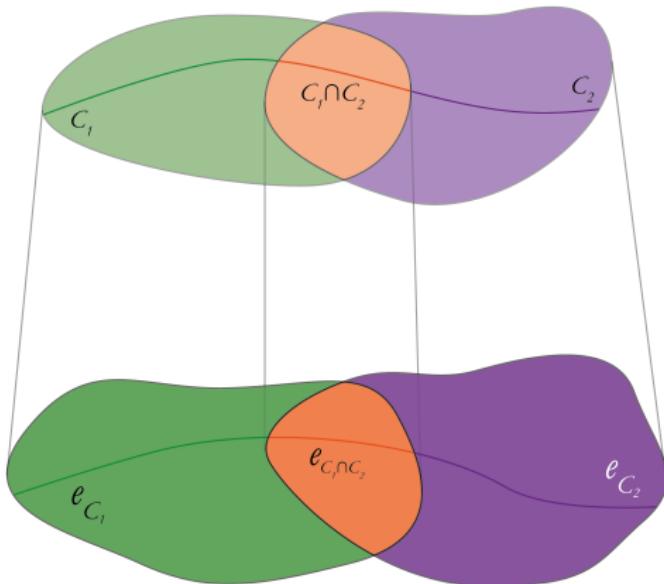
A crash course in sheaf theory

Sheaves



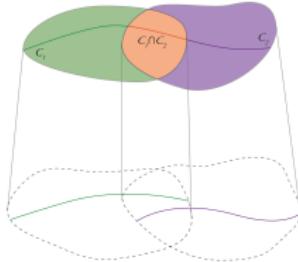
A crash course in sheaf theory

Sheaves



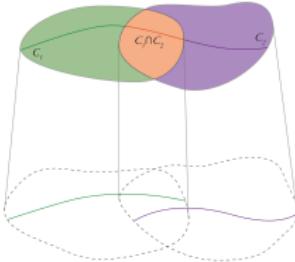
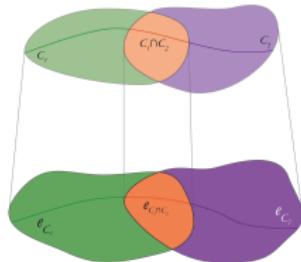
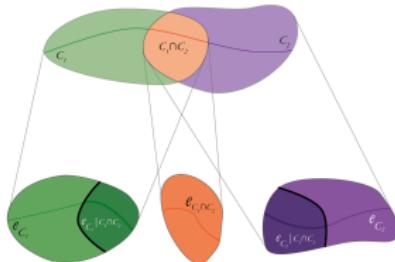
A crash course in sheaf theory

SF: the signalling-fraction



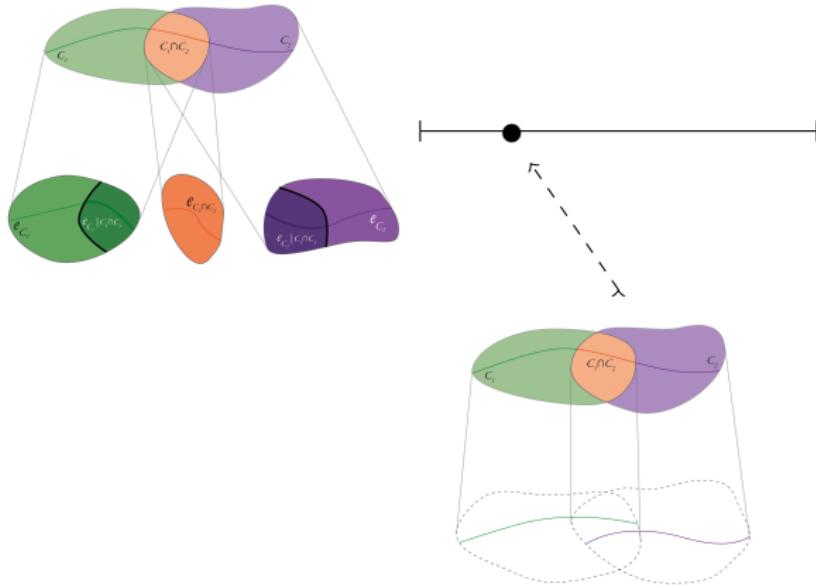
A crash course in sheaf theory

SF: the signalling-fraction



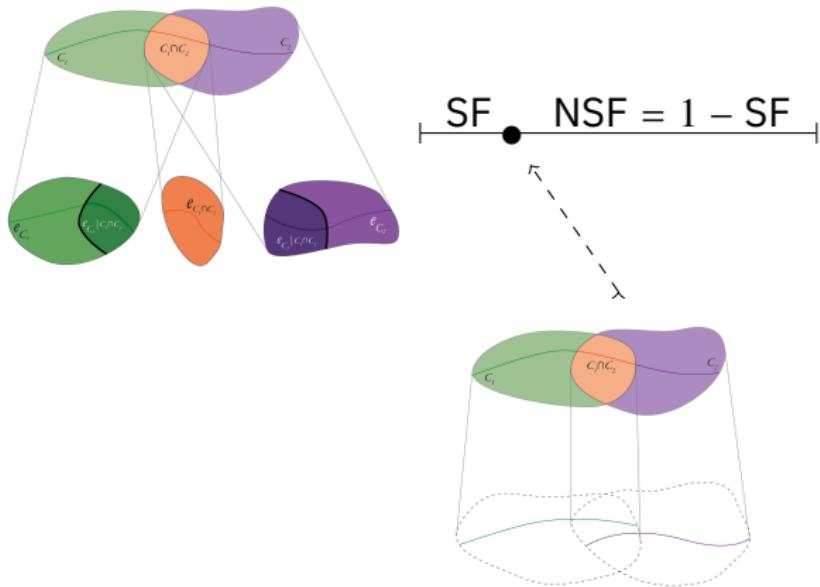
A crash course in sheaf theory

SF: the signalling-fraction



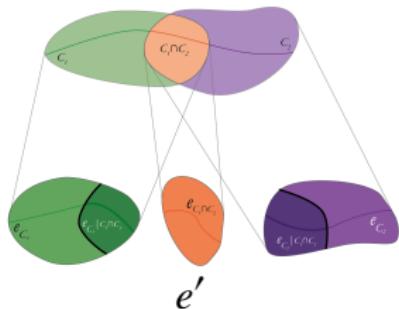
A crash course in sheaf theory

SF: the signalling-fraction



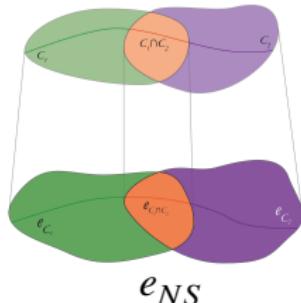
A crash course in sheaf theory

SF: the signalling-fraction



$$\text{SF} \quad \bullet \quad \text{NSF} = 1 - \text{SF}$$

$\nwarrow \lambda$



A diagram showing a presheaf $e = \text{SF}e' + \text{NSF } e_{NS}$ over a topological space $C_i \cap C_j$. The space is represented by a green oval C_i and a purple oval C_j intersecting at a central orange region $C_i \cap C_j$. Three arrows point from three separate regions labeled $\ell_{C_i}|_{C_i \cap C_j}$, $\ell_{C_j}|_{C_i \cap C_j}$, and $\ell_{C_i \cap C_j}$ to the central intersection point. Dashed lines indicate the continuation of the presheaf components beyond the intersection.

The model

The context space

g =The employees understood...

The model

The context space

$g = \text{The employees understood...} \longrightarrow \Sigma(g) = \{\text{the, employees, understood, contract ...}\}$

The model

The context space

$g = \text{The employees understood...} \longrightarrow \Sigma(g) = \{\text{the, employees, understood, contract ...}\}$

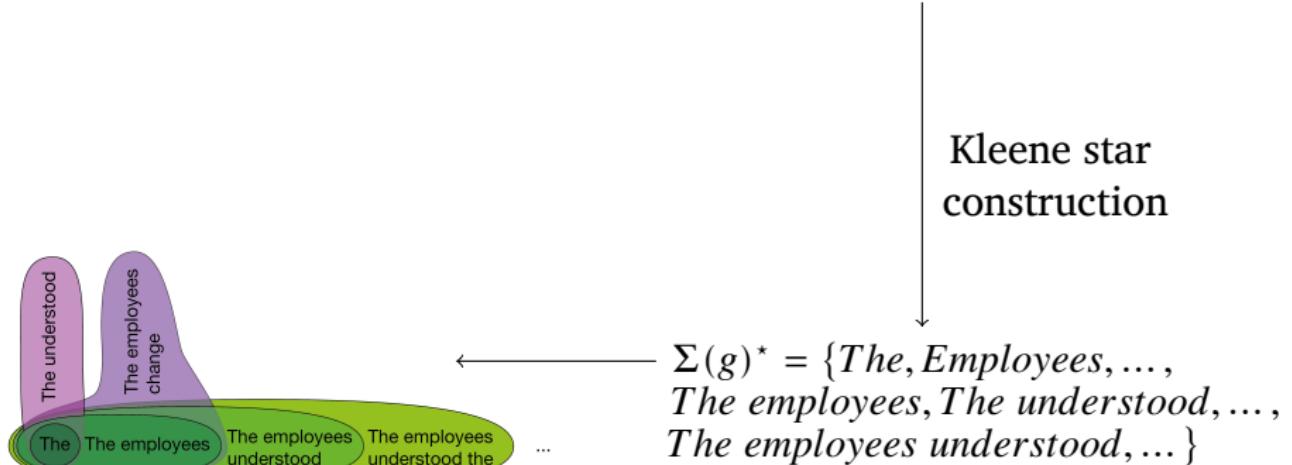
Kleene star
construction

$\Sigma(g)^* = \{\text{The, Employees, ...},$
 $\text{The employees, The understood, ...},$
 $\text{The employees understood, ...}\}$

The model

The context space

$g = \text{The employees understood...} \longrightarrow \Sigma(g) = \{\text{the, employees, understood, contract ...}\}$



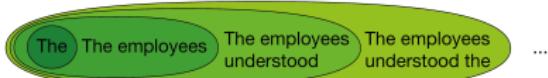
The model

The context space

$g = \text{The employees understood...} \longrightarrow \Sigma(g) = \{\text{the, employees, understood, contract ...}\}$

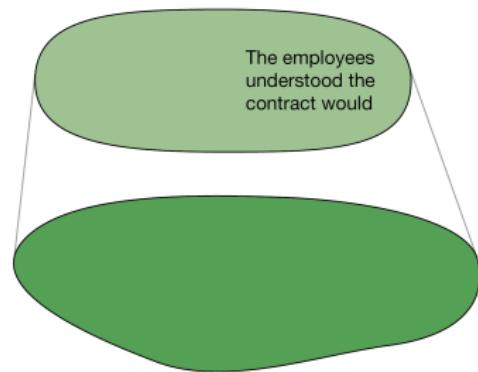
Kleene star construction

$\Sigma(g)^* = \{\text{The, Employees, ...}, \text{The employees, The understood, ...}, \text{The employees understood, ...}\}$



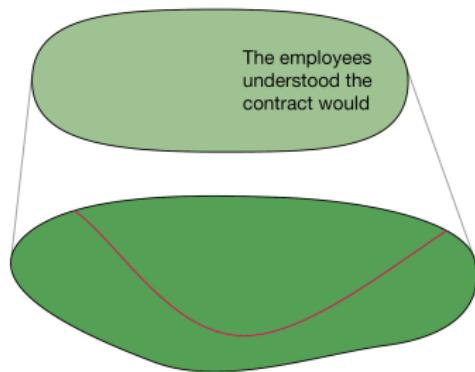
The model

The data space



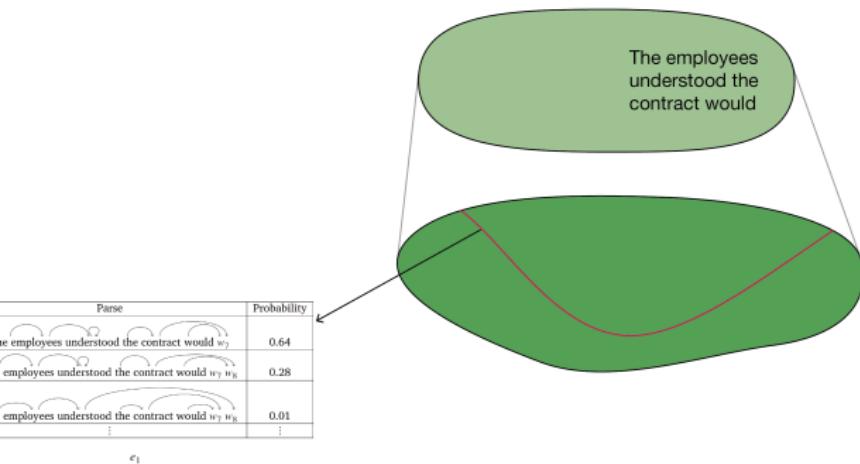
The model

The data space



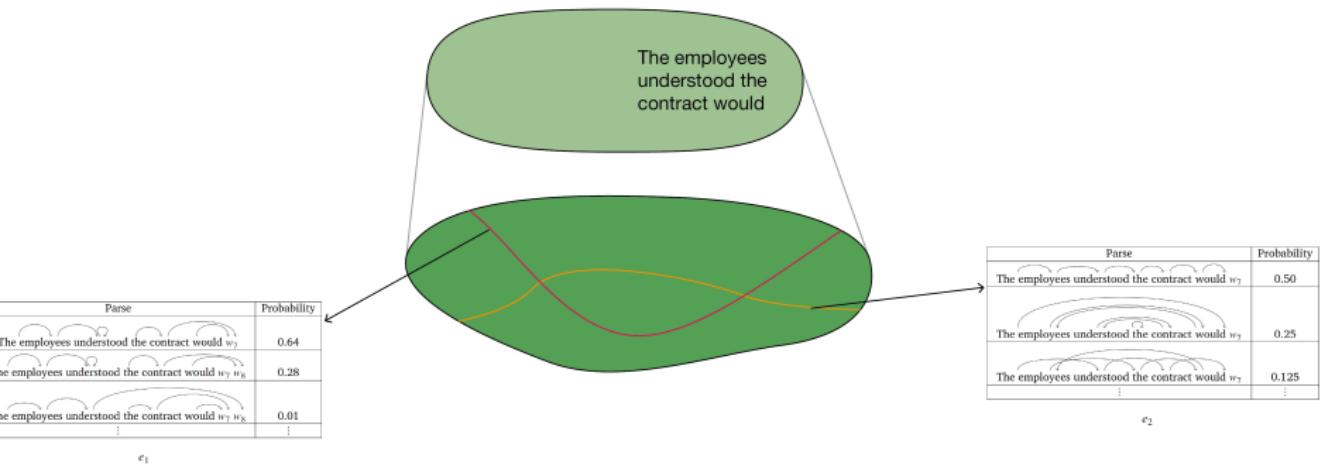
The model

The data space



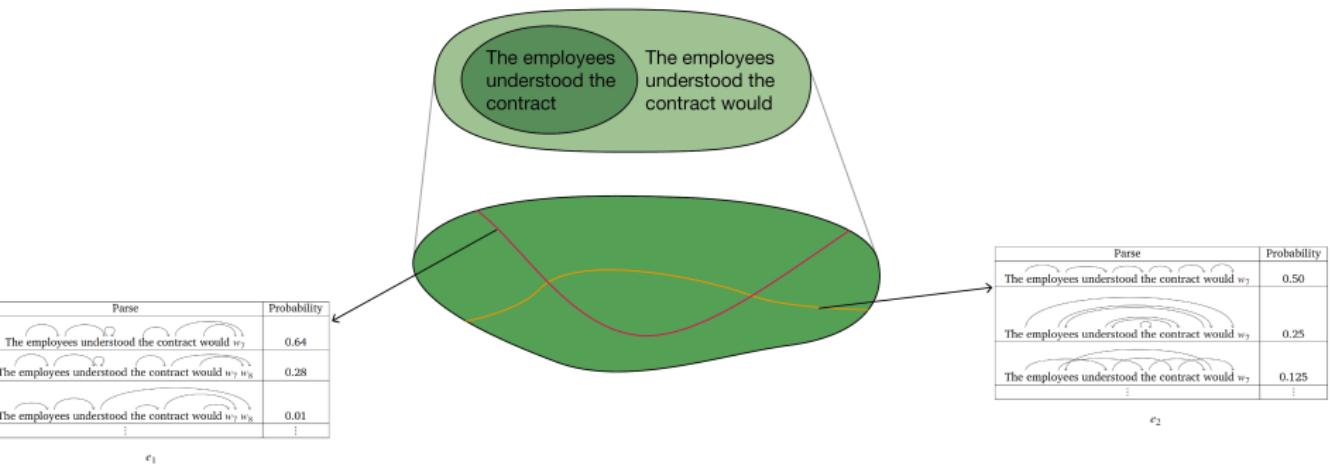
The model

The data space



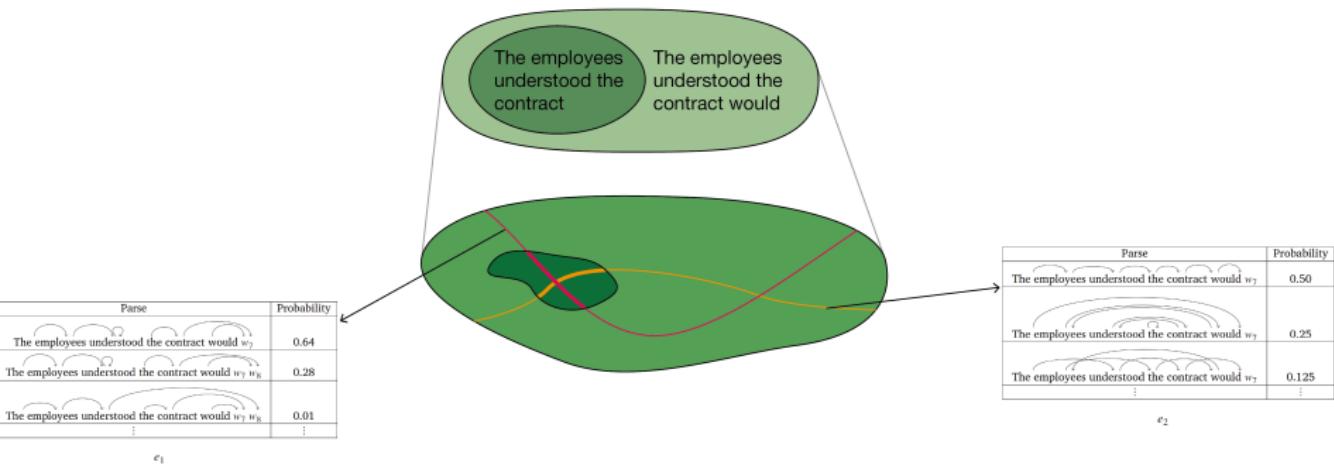
The model

The data space



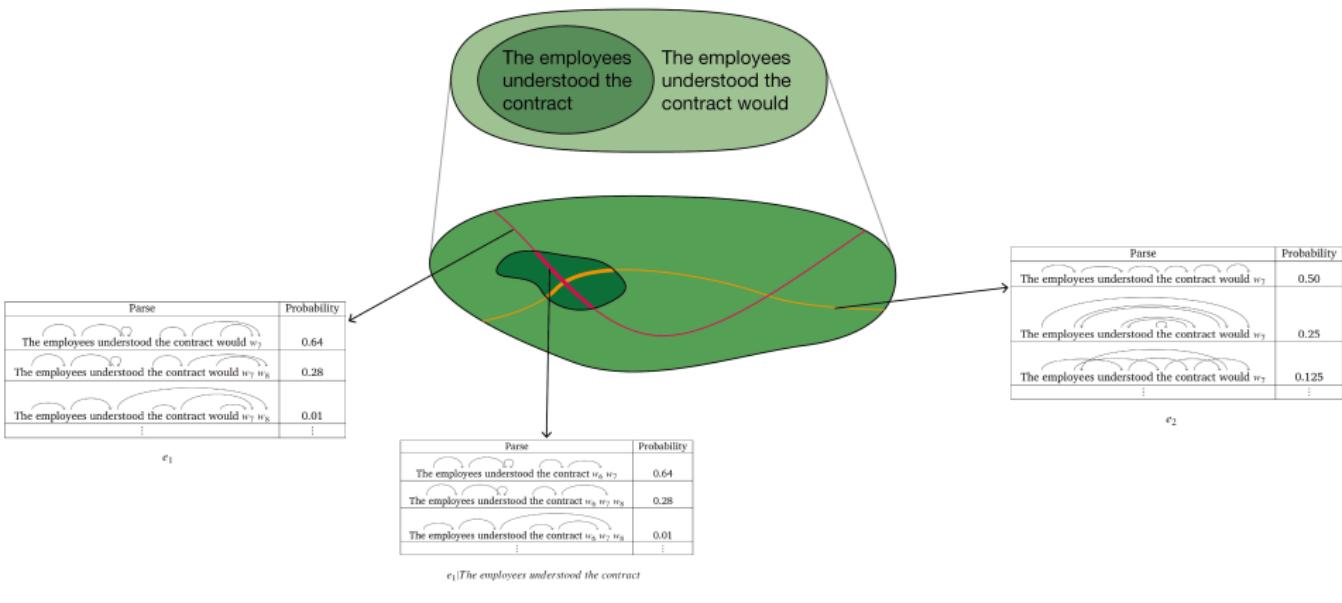
The model

The data space



The model

The data space



The model

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮

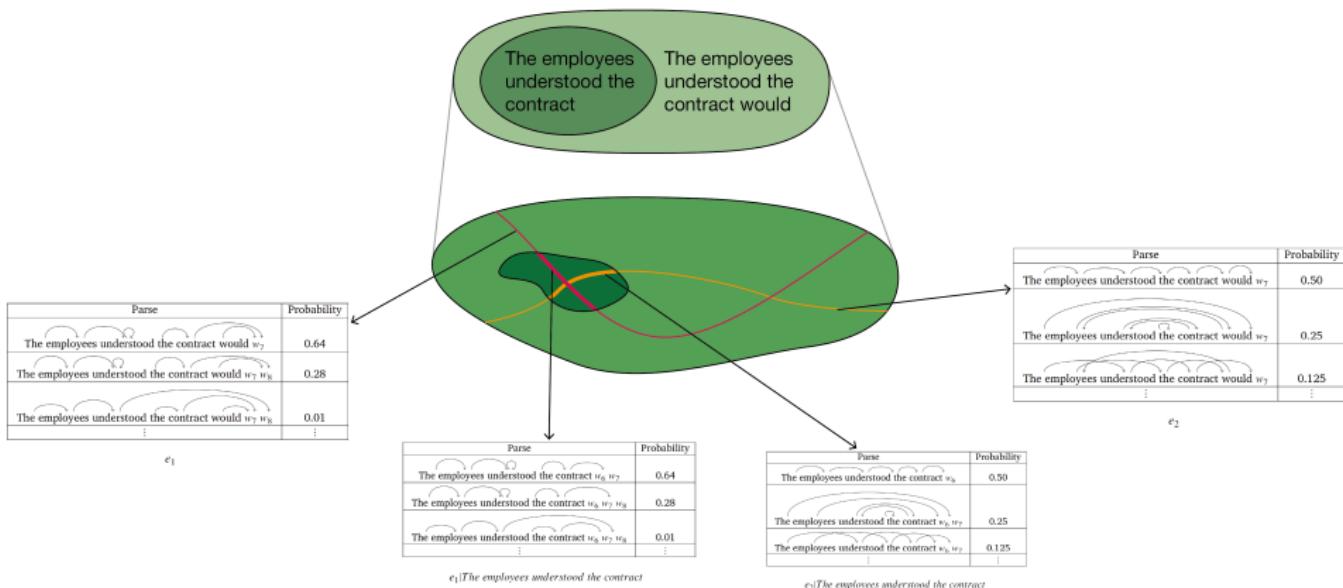
 e_1 

Parse	Probability
The employees understood the contract $w_6 w_7$	0.64
The employees understood the contract $w_6 w_7 w_8$	0.28
The employees understood the contract $w_6 w_7 w_8$	0.01
⋮	⋮

 $e_1 | \text{The employees understood the contract}$

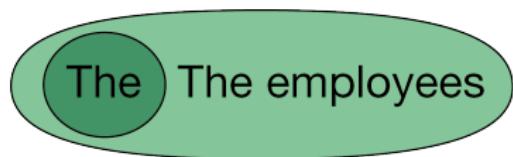
The model

The data space



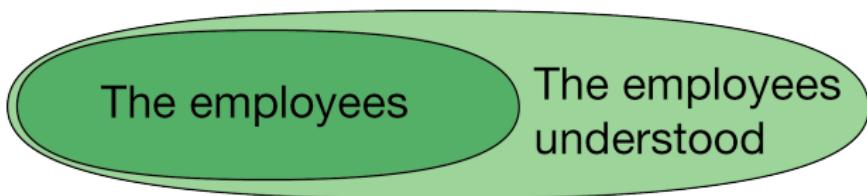
The model

Incrementality: the contexts



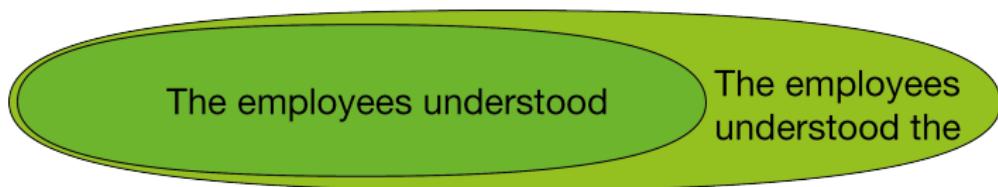
The model

Incrementality: the contexts



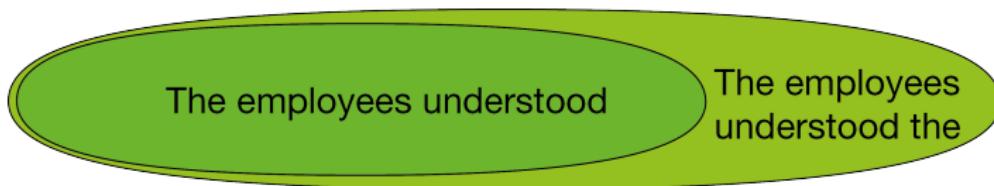
The model

Incrementality: the contexts



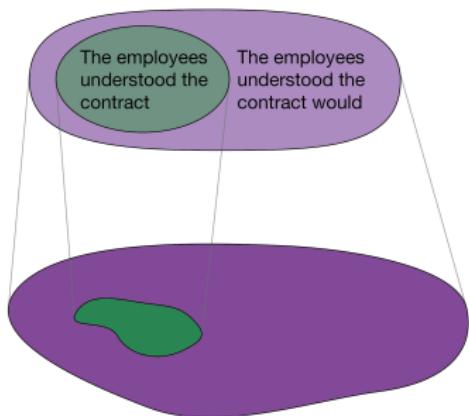
The model

Incrementality: the contexts



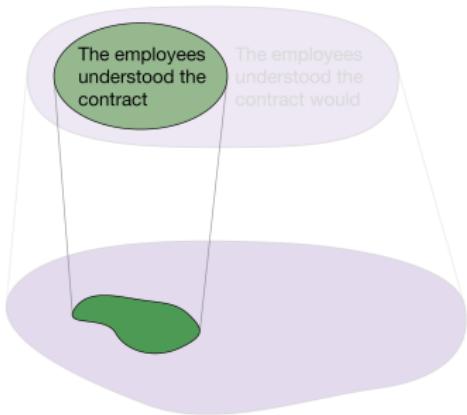
The model

Incrementatily: The data & SF



The model

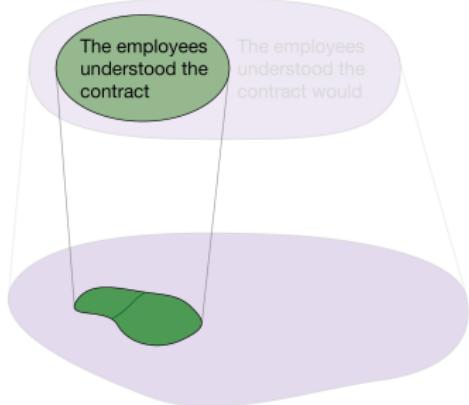
Incrementatily: The data & SF



The model

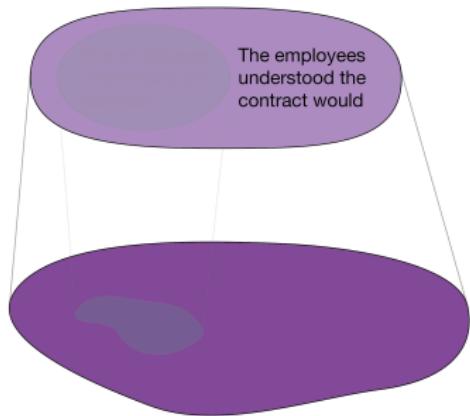
Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮



The model

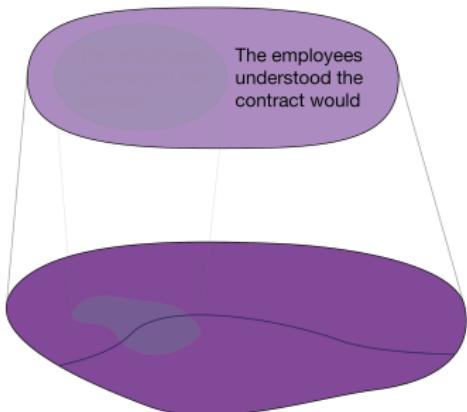
Incrementatily: The data & SF



The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮

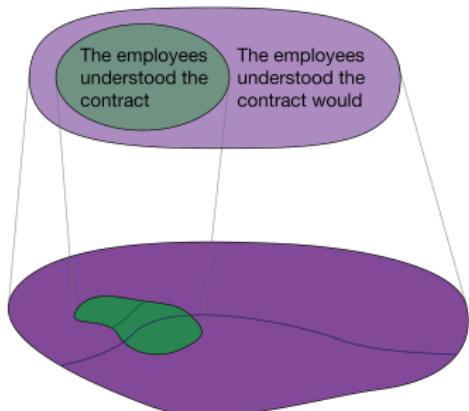


The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮

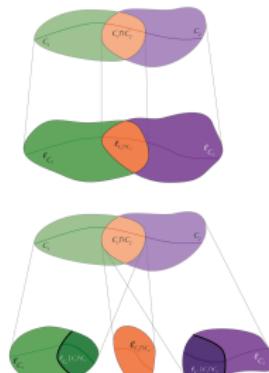


The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮

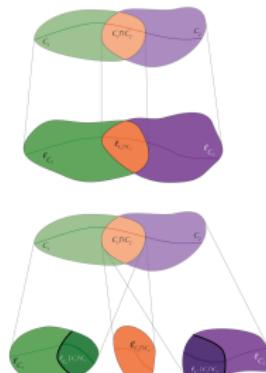


The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮



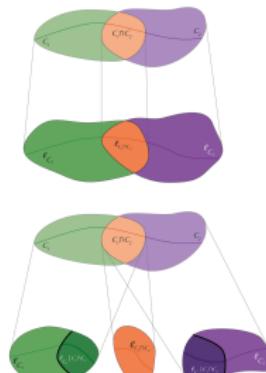
No need for reanalysis

The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮



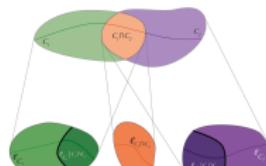
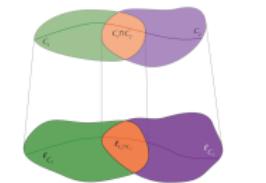
No need for
reanalysis
(EASY)

The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮



No need for
reanalysis
(EASY)

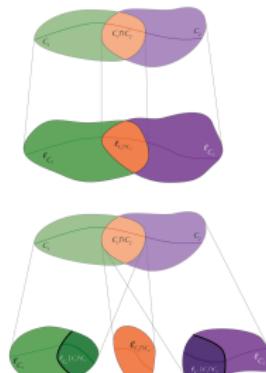
Need to
reanalyse

The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮



No need for
reanalysis
(EASY)

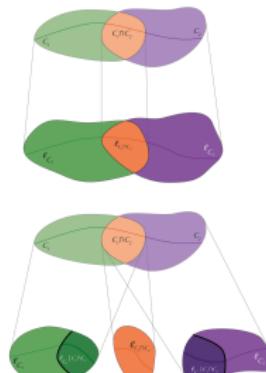
Need to
reanalyse
(HARD)

The model

Incrementatily: The data & SF

Parse	Probability
The employees understood the contract	0.37
The employees understood the contract w_6	0.24
The employees understood the contract $w_6 w_7$	0.15
⋮	⋮

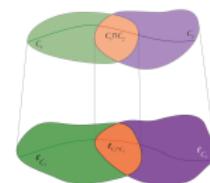
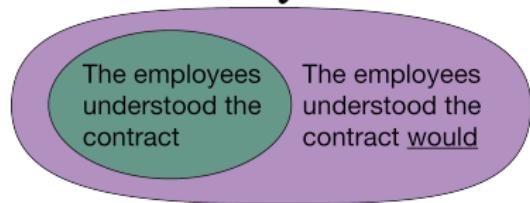
Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮



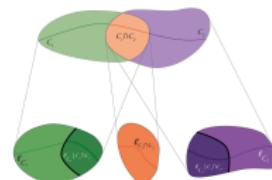
No need for
reanalysis
(EASY)

Need to
reanalyse
(HARD)

The model

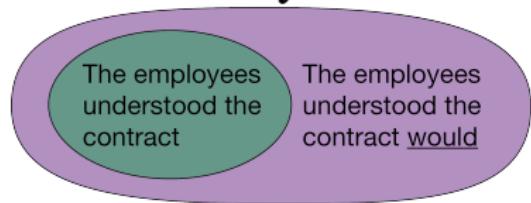
Incrementatily: The data & SF

No need for
reanalysis
(EASY)

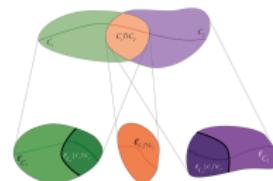
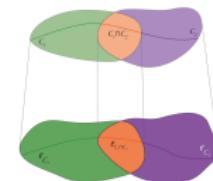


Need to
reanalyse
(HARD)

The model

Incrementatily: The data & SF

SF should quantify how much the reader will have to change its grammatical analysis



No need for reanalysis
(EASY)

Need to reanalyse
(HARD)

The model

Examples

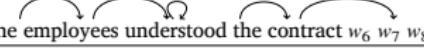
Extreme cases: SF = 0 (completely consistent)

The model

Examples

Extreme cases: SF = 0 (completely consistent)

*“The employees understood
the contract”*

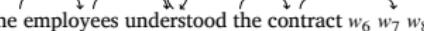
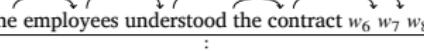
Parse	Probability
The employees understood the contract  $w_6 \ w_7$	0.64
The employees understood the contract  $w_6 \ w_7 \ w_8$	0.28
The employees understood the contract  $w_6 \ w_7 \ w_8$	0.01
⋮	⋮

The model

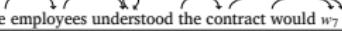
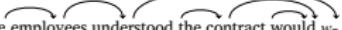
Examples

Extreme cases: SF = 0 (completely consistent)

*“The employees understood
the contract”*

Parse	Probability
 The employees understood the contract $w_6 \ w_7$	0.64
 The employees understood the contract $w_6 \ w_7 \ w_8$	0.28
 The employees understood the contract $w_6 \ w_7 \ w_8$	0.01
⋮	⋮

*“The employees understood
the contract would”*

Parse	Probability
 The employees understood the contract would w_7	0.64
 The employees understood the contract would $w_7 \ w_8$	0.28
 The employees understood the contract would $w_7 \ w_8$	0.01
⋮	⋮

The model

Examples

Extreme cases: SF = 0 (completely consistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	0.64
The employees understood the contract $w_6 w_7 w_8$	0.28
The employees understood the contract $w_6 w_7 w_8$	0.01
⋮	⋮

*“The employees understood
the contract would”*

Parse	Probability
The employees understood the contract would w_7	0.64
The employees understood the contract would $w_7 w_8$	0.28
The employees understood the contract would $w_7 w_8$	0.01
⋮	⋮

Parse	Probability
The employees understood the contract $w_6 w_7$	0.64
The employees understood the contract $w_6 w_7 w_8$	0.28
The employees understood the contract $w_6 w_7 w_8$	0.01
⋮	⋮

The model

Examples

Extreme cases: $SF = 1$ (completely inconsistent)

The model

Examples

Extreme cases: SF = 1 (completely inconsistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	1
other parses	0

The model

Examples

Extreme cases: SF = 1 (completely inconsistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	1
other parses	0

*“The employees understood
the contract would”*

Parse	Probability
The employees understood the contract would w_7	1
other parses	0

The model

Examples

Extreme cases: SF = 1 (completely inconsistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	1
other parses	0

*“The employees understood
the contract would”*

Parse	Probability
The employees understood the contract would w_7	1
other parses	0

↓

Parse	Probability
The employees understood the contract w_6	1
other parses	0

The model

Examples

Middle cases: $SF \simeq 0$ (mostly consistent)

The model

Examples

Middle cases: $SF \approx 0$ (mostly consistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	0.8
The employees understood the contract w_6 other parses	0.2
	0

The model

Examples

Middle cases: $SF \approx 0$ (mostly consistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	0.8
The employees understood the contract w_6	0.2
other parses	0

*“The employees understood
the contract would”*

Parse	Probability
The employees understood the contract would w_7	0.9
The employees understood the contract would w_7 other parses	0.1
other parses	0

The model

Examples

Middle cases: $SF \approx 0$ (mostly consistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	0.8
The employees understood the contract w_6	0.2
other parses	0

*“The employees understood
the contract would”*

Parse	Probability
The employees understood the contract would w_7	0.9
The employees understood the contract would w_7	0.1
other parses	0

Parse	Probability
The employees understood the contract $w_6 w_7$	0.9
The employees understood the contract w_6	0.1
other parses	0

The model

Examples

Middle cases: $SF \approx 1$ (mostly inconsistent)

The model

Examples

Middle cases: $SF \approx 1$ (mostly inconsistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	0.8
The employees understood the contract w_6 other parses	0.2
	0

The model

Examples

Middle cases: $SF \approx 1$ (mostly inconsistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	0.8
The employees understood the contract w_6	0.2
other parses	0

*“The employees understood
the contract would”*

Parse	Probability
The employees understood the contract would w_7	0.1
The employees understood the contract would w_7 other parses	0.9
	0

The model

Examples

Middle cases: $SF \approx 1$ (mostly inconsistent)

*“The employees understood
the contract”*

Parse	Probability
The employees understood the contract $w_6 w_7$	0.8
The employees understood the contract w_6	0.2
other parses	0

*“The employees understood
the contract would”*

Parse	Probability
The employees understood the contract would w_7	0.1
The employees understood the contract would w_7	0.9
other parses	0

Parse	Probability
The employees understood the contract $w_6 w_7$	0.1
The employees understood the contract w_6	0.9
other parses	0

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
⋮	⋮

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



The employees understood the need for ...



Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



The employees understood the need for ...



spaCy



Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

The employees understood the need for ...



spaCy



The employees understood the need for ...

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Full parse	Probability
The employees understood the need for ...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks...	0.015
⋮	⋮

Prediction	Probability
The employees understood the need for ...	0.053
The employees understood the seriousness of ...	0.021
The employees understood the security risks ...	0.015
⋮	⋮

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

The employees understood the need for ...



spaCy



The employees understood the need for ...

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

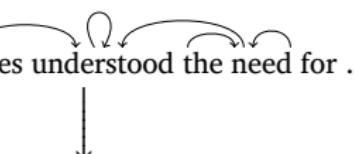
The employees understood the need for ...



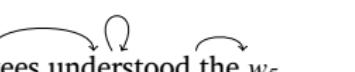
spaCy



The employees understood the need for ...



The employees understood the w_5 ...



Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Partial parse	Probability
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $\overset{\curvearrowleft}{w_5}$...	0.053
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $\overset{\curvearrowleft}{w_5}$ $\overset{\curvearrowleft}{w_6}$...	0.021
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $w_5 \overset{\curvearrowleft}{w_6}$...	0.015
:	:

Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Partial parse	Probability
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $\overset{\curvearrowleft}{w_5}$...	0.053
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $\overset{\curvearrowleft}{w_5}$ $\overset{\curvearrowleft}{w_6}$...	0.021
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $w_5 \overset{\curvearrowleft}{w_6}$...	0.015
:	:

Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
:	:

Collecting data using LLMs

The employees understood the [MASK] [MASK] ...



Partial parse	Probability
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $\overset{\curvearrowleft}{w_5}$...	0.074
The $\overset{\curvearrowleft}{\text{employees}}$ $\overset{\curvearrowleft}{\text{understood}}$ the $w_5 \overset{\curvearrowleft}{w_6}$...	0.015
⋮	⋮

Prediction	Probability
The employees understood the need for...	0.053
The employees understood the seriousness of...	0.021
The employees understood the security risks ...	0.015
⋮	⋮



Data analysis

The Sturt/Pickering/Crocker dataset

Description of the dataset

- ▶ 32 NP/S sentences (+ unambiguous version) & 32 NP/Z sentences (+ unambiguous version)

Sturt, Patrick, Martin J. Pickering, and Matthew W. Crocker. "Structural change and reanalysis difficulty in language comprehension." *Journal of Memory and Language* 40.1 (1999): 136-150.

The Sturt/Pickering/Crocker dataset

Description of the dataset

- ▶ 32 NP/S sentences (+ unambiguous version) & 32 NP/Z sentences (+ unambiguous version)
- ▶ NP/S and NP/Z sentences with overlapping vocabulary

Sturt, Patrick, Martin J. Pickering, and Matthew W. Crocker. "Structural change and reanalysis difficulty in language comprehension." *Journal of Memory and Language* 40.1 (1999): 136-150.

The Sturt/Pickering/Crocker dataset

Description of the dataset

- ▶ 32 NP/S sentences (+ unambiguous version) & 32 NP/Z sentences (+ unambiguous version)
- ▶ NP/S and NP/Z sentences with overlapping vocabulary
- ▶ Each sentence is split into regions

Sturt, Patrick, Martin J. Pickering, and Matthew W. Crocker. "Structural change and reanalysis difficulty in language comprehension." *Journal of Memory and Language* 40.1 (1999): 136-150.

The Sturt/Pickering/Crocker dataset

Description of the dataset

- ▶ 32 NP/S sentences (+ unambiguous version) & 32 NP/Z sentences (+ unambiguous version)
- ▶ NP/S and NP/Z sentences with overlapping vocabulary
- ▶ Each sentence is split into regions

Regions				
NP/S (ambiguous)	The faithful employees	<u>understood the technical contract</u>	would be changed	very soon
NP/S (unambiguous)	<u>The faithful employees</u>	<u>understood that the technical contract</u>	would be changed	every soon
NP/Z (ambiguous)	Because the employees	<u>negotiated the technical contract</u>	would be changed	every soon
NP/Z (unambiguous)	<u>Because the employees</u>	<u>negotiated, the technical contract</u>	would be changed	every soon

Sturt, Patrick, Martin J. Pickering, and Matthew W. Crocker. "Structural change and reanalysis difficulty in language comprehension." *Journal of Memory and Language* 40.1 (1999): 136-150.

The Sturt/Pickering/Crocker dataset

Description of the dataset

- ▶ 32 NP/S sentences (+ unambiguous version) & 32 NP/Z sentences (+ unambiguous version)
- ▶ NP/S and NP/Z sentences with overlapping vocabulary
- ▶ Each sentence is split into regions

Regions				
NP/S (ambiguous)	The faithful employees	<u>understood the technical contract</u>	would be changed	very soon
NP/S (unambiguous)	The faithful employees	<u>understood that the technical contract</u>	would be changed	every soon
NP/Z (ambiguous)	Because the employees	<u>negotiated the technical contract</u>	would be changed	every soon
NP/Z (unambiguous)	Because the employees	<u>negotiated, the technical contract</u>	would be changed	every soon

- ▶ Self-paced reading (region by region)

Sturt, Patrick, Martin J. Pickering, and Matthew W. Crocker. "Structural change and reanalysis difficulty in language comprehension." *Journal of Memory and Language* 40.1 (1999): 136-150.

The Sturt/Pickering/Crocker dataset

Description of the dataset

- ▶ 32 NP/S sentences (+ unambiguous version) & 32 NP/Z sentences (+ unambiguous version)
- ▶ NP/S and NP/Z sentences with overlapping vocabulary
- ▶ Each sentence is split into regions

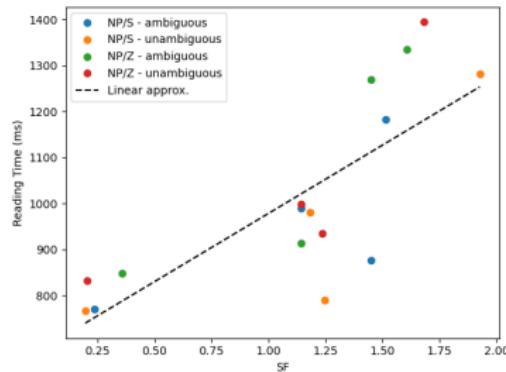
Regions				
NP/S (ambiguous)	The faithful employees	<u>understood the technical contract</u>	would be changed	very soon
NP/S (unambiguous)	The faithful employees	<u>understood that the technical contract</u>	would be changed	every soon
NP/Z (ambiguous)	Because the employees	<u>negotiated the technical contract</u>	would be changed	every soon
NP/Z (unambiguous)	Because the employees	<u>negotiated, the technical contract</u>	would be changed	every soon

- ▶ Self-paced reading (region by region)
- ▶ Averaged reading time for each region, for each type of sentence (e.g. NP/S ambiguous, NP/S unambiguous, ...)

Sturt, Patrick, Martin J. Pickering, and Matthew W. Crocker. "Structural change and reanalysis difficulty in language comprehension." *Journal of Memory and Language* 40.1 (1999): 136-150.

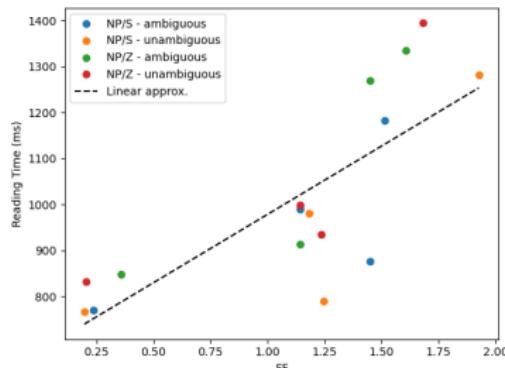
**Does SF correlate with
reading times?**

Linear Correlations



$$RT(R) = 297 \times \sum_{w \in R} SF(w) + 682 \text{ ms}$$

Linear Correlations



$$RT(R) = 297 \times \sum_{w \in R} SF(w) + 682 \text{ ms}$$

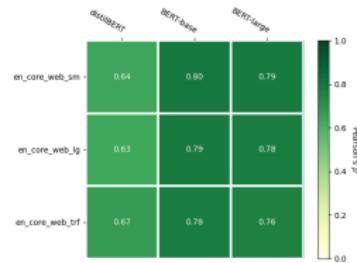


Figure: Pearson's ρ -coefficients

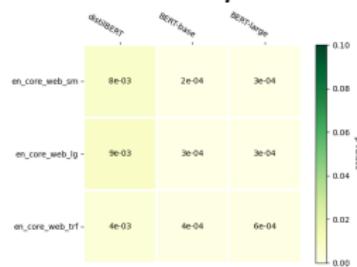


Figure: p -values associated with the Pearson's ρ

Detecting a garden-path effect

**Do we predict higher
reading times for
garden-path sentences?**

Detecting a garden-path effect

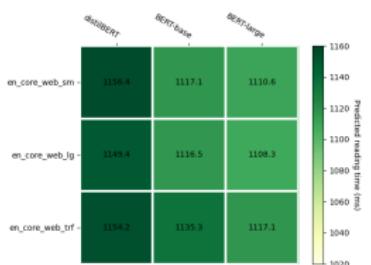


Figure: Reading times for ambiguous sentences

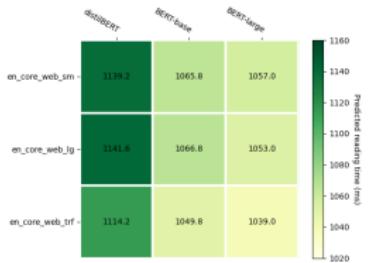


Figure: Reading times for unambiguous sentences

Detecting a garden-path effect

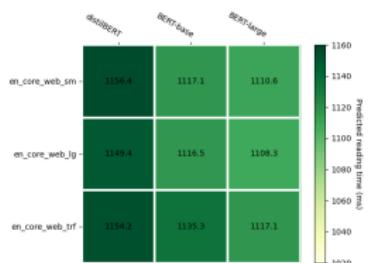


Figure: Reading times for ambiguous sentences

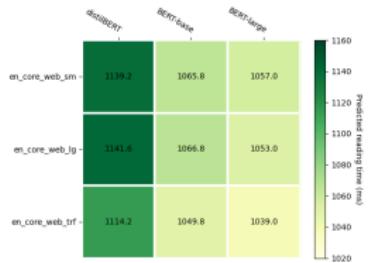


Figure: Reading times for unambiguous sentences

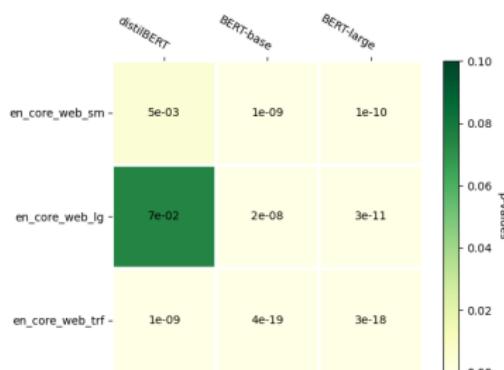


Figure: p -values for the 1-sample t -test (reading times for the garden-path sentence vs. their unambiguous equivalent)

**Can we see that NP/Z
sentences are harder to
read than NP/S
sentences?**

Effects for NP/S and NP/Z sentences

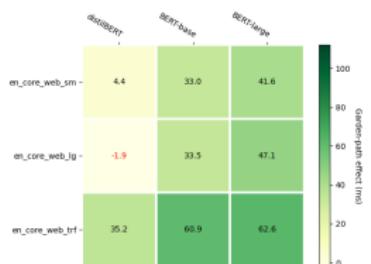


Figure: Garden-path effect for NP/S sentences

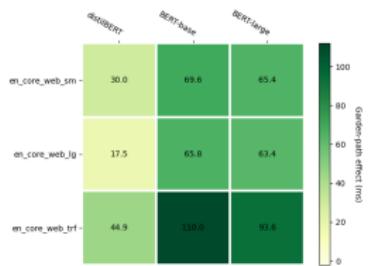


Figure: Garden-path effect for NP/Z sentences

Effects for NP/S and NP/Z sentences

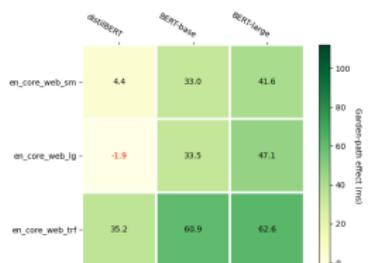


Figure: Garden-path effect for NP/S sentences

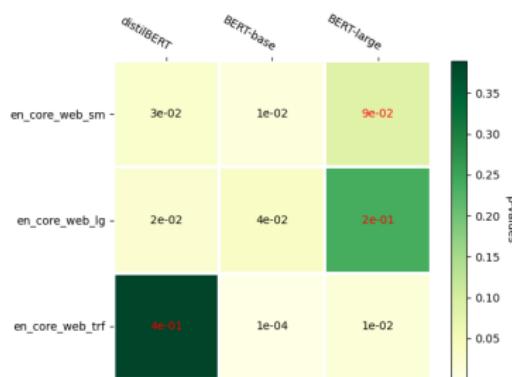


Figure: p -values for t -test (garden-path effects for NP/S vs. NP/Z)

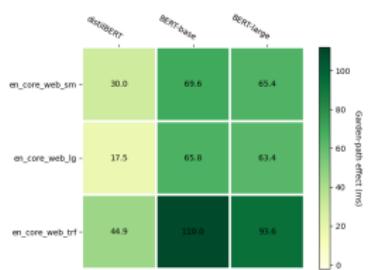
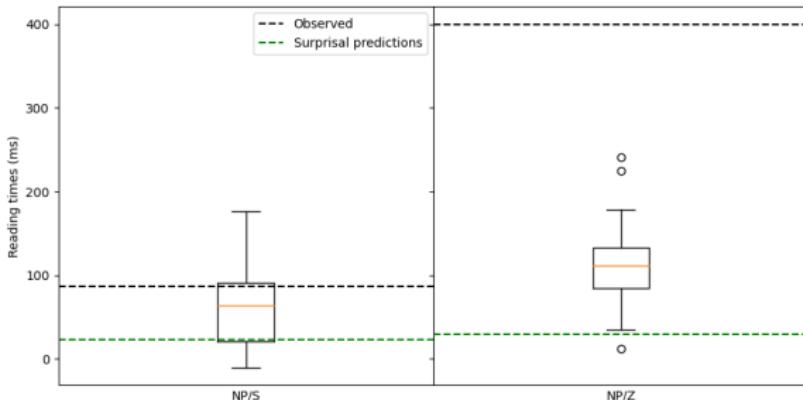


Figure: Garden-path effect for NP/Z sentences

**Can we do better than
surprisal?**

Comparison with surprisal predictions



	Prediction (ms)		Observed (ms)
	SF	S^1	
NP/S	63	24	87
NP/Z	110	30	400

¹van Schijndel, Marten, and Tal Linzen. “Modeling garden path effects without explicit hierarchical syntax.” *CogSci*. 2018.

Conclusion

Summary of the results

- ▶ Created a presheaf model of the human parsing process, which is close to the theories of psycholinguistics

Summary of the results

- ▶ Created a presheaf model of the human parsing process, which is close to the theories of psycholinguistics
- ▶ We used a measure of inconsistency SF to quantify the difficulty of parsing words in a sentence

Summary of the results

- ▶ Created a presheaf model of the human parsing process, which is close to the theories of psycholinguistics
- ▶ We used a measure of inconsistency SF to quantify the difficulty of parsing words in a sentence
- ▶ We obtained good correlations between SF and reading times

Summary of the results

- ▶ Created a presheaf model of the human parsing process, which is close to the theories of psycholinguistics
- ▶ We used a measure of inconsistency SF to quantify the difficulty of parsing words in a sentence
- ▶ We obtained good correlations between SF and reading times
- ▶ We managed to obtain statistically differences between predictions from garden-path sentences which have different levels of difficulty

Summary of the results

- ▶ Created a presheaf model of the human parsing process, which is close to the theories of psycholinguistics
- ▶ We used a measure of inconsistency SF to quantify the difficulty of parsing words in a sentence
- ▶ We obtained good correlations between SF and reading times
- ▶ We managed to obtain statistically differences between predictions from garden-path sentences which have different levels of difficulty
- ▶ We compared our results with the state-of-the-art methods from computational linguistics, and obtained more accurate predictions

Future work

- ▶ Investigate the link between SF and surprisal

Future work

- ▶ Investigate the link between SF and surprisal
- ▶ Make use of more detailed datasets

Future work

- ▶ Investigate the link between SF and surprisal
- ▶ Make use of more detailed datasets
- ▶ Compare the results with other possible measures that can arise from this framework

Future work

- ▶ Investigate the link between SF and surprisal
- ▶ Make use of more detailed datasets
- ▶ Compare the results with other possible measures that can arise from this framework
- ▶ Use a variation of this sheaf-theoretic framework to study reanalysis patterns

Future work

- ▶ Investigate the link between SF and surprisal
- ▶ Make use of more detailed datasets
- ▶ Compare the results with other possible measures that can arise from this framework
- ▶ Use a variation of this sheaf-theoretic framework to study reanalysis patterns
- ▶ Apply this framework to other phenomena (e.g. recursive structures, memory effect, ...)

Thank you!

References

Our work

- ▶ Wang D., Sadrzadeh M. “Causality and Signalling of Garden-Path Sentences”. To appear in *Philosophical Transactions of Royal Society A*
- ▶ Dataset: github.com/wangdaphne/garden-path-SF-dataset

Surprisal in garden-path sentences

- ▶ Hale, John. “A probabilistic Earley parser as a psycholinguistic model.” *Second meeting of the north american chapter of the association for computational linguistics*. 2001.
- ▶ van Schijndel, Marten, and Tal Linzen. “Modeling garden path effects without explicit hierarchical syntax.” *CogSci*. 2018.
- ▶ van Schijndel, Marten, and Tal Linzen. “Single-stage prediction models do not explain the magnitude of syntactic disambiguation difficulty.” *Cognitive science* 45.6 (2021).

References (cont.)

- ▶ Arehalli, Suhas, Brian Dillon, and Tal Linzen. “Syntactic surprisal from neural models predicts, but underestimates, human processing difficulty from syntactic ambiguities.” *arXiv:2210.12187* (2022).

Psycholinguistics of garden-path sentences

- ▶ Bever, Thomas G. “The cognitive basis for linguistic structures.” *Cognition and the development of language* (1970).
- ▶ Frazier, Lyn. *On comprehending sentences: Syntactic parsing strategies*. University of Connecticut, 1979.
- ▶ van Gompel, Roger PG, et al. “Evidence against competition during syntactic ambiguity resolution.” *Journal of Memory and Language* 52.2 (2005): 284-307.

Sheaf theory

- ▶ MacLane, Saunders, and Ieke Moerdijk. *Sheaves in geometry and logic: A first introduction to topos theory*. 2012.