Towards a GLUCOSE Knowledge Graph

and the power of it's generalization

Nadia Arslan

University of Heidelberg
Institute of Computational Linguistics
BA-Thesis
Prof. Dr. Anette Frank

01.07.2021

Table of Contents

What is GLUCOSE?

GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* **Aditya Kalyanpur** **Lori Moon** **David Buchanan†**
Lauren Berkowitz **Or Biran** **Jennifer Chu-Carroll**
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an inference rule generalized from the statement. This paper details two concrete contributions. First, we present our platform for effectively crowdsourcing GLUCOSE data at scale, which uses semi-structured templates to elicit causal explanations. Using this platform, we collected a total of ~670K specific statements and general rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

[1] Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

What is GLUCOSE?

GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* Aditya Kalyanpur Lori Moon David Buchanan†
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as mini-theories about the world, each grounded in a narrative context.

eral rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent ... happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). ... they can construct the causal chain that explains how the car's unexpected turn ultimately ... falling, describe how Gage's emotion and location changed throughout the story, ... that he likely shouted for help ...

Though humans build such mental models with ... (... et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far ... similar commonsense reasoning ... Two major bottlenecks have been acquiring commonsense knowledge and successfully ... into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

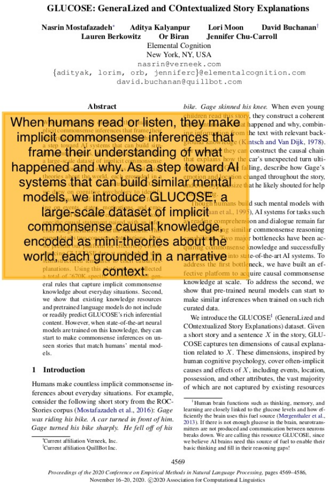
[1] Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

What is GLUCOSE?


GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* Aditya Kalyanpur Lori Moon David Buchanan†
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as mini-theories about the world, each grounded in a narrative context.

eral rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent ...

address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

[1] Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586, November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations: (Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge

What is GLUCOSE?

GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh[*] Aditya Kalyanpur Lori Moon David Buchanan[†]
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an inference rule generalized from the statement. This paper details two concrete contributions. First, we present our platform for effectively crowdsourcing GLUCOSE data at scale, which uses semi-structured templates to elicit causal explanations. Using this platform, we collected a total of ˜670K specific statements and general rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense in- ... consider the following short story from the ROC- ... *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

[*]Current affiliation Verneek, Inc.
[†]Current affiliation QuillBot Inc.

[1]Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

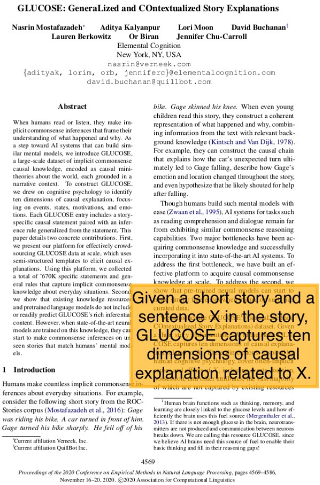
4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)

What is GLUCOSE?



GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh[*] Aditya Kalyanpur Lori Moon David Buchanan[†]
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an inference rule generalized from the statement. This paper details two concrete contributions. First, we present our platform for effectively crowdsourcing GLUCOSE data at scale, which uses semi-structured templates to elicit causal explanations. Using this platform, we collected a total of ~670K specific statements and general rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to [illegible] curated data. [illegible] COntextualized Story Explanations) dataset. Given [illegible] COSE captures ten dimensions of causal explana- [illegible] human cognitive psychology, cover often-implicit [illegible] of which are not captured by existing resources

[*]Current affiliation Verneek, Inc.
[†]Current affiliation QuillBot Inc.

[1]Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

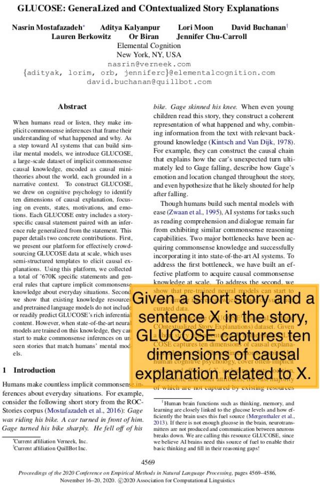
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X.

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)

What is GLUCOSE?



GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* Aditya Kalyanpur Lori Moon David Buchanan†
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal minitheories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an inference rule generalized from the statement. This paper details two concrete contributions. First, we present our platform for effectively crowdsourcing GLUCOSE data at scale, which uses semi-structured templates to elicit causal explanations. Using this platform, we collected a total of ~670K specific statements and general rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to ... curated data.

... COntextualized Story Explanations) dataset. Given a short story ... COSE captures ten dimensions of causal explana... human cognitive psychology, cover often-implicit ... possession, and other attributes, the vast majority of which are not captured by existing resources

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

[1] Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

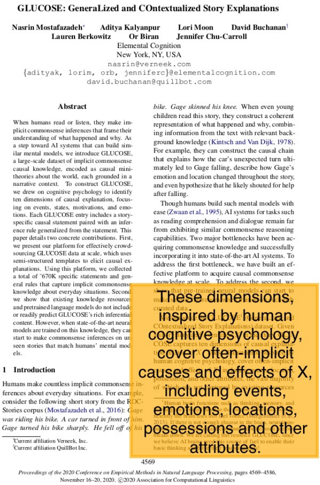
4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)
- captures different dimensions of causal explanation

What is GLUCOSE?



GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* Aditya Kalyanpur Lori Moon David Buchanan†
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an inference rule generalized from the statement. This paper details two concrete contributions. First, we present our platform for effectively crowdsourcing GLUCOSE data at scale, which uses semi-structured templates to elicit causal explanations. Using this platform, we collected a total of ~670K specific statements and general rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we …

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

4569

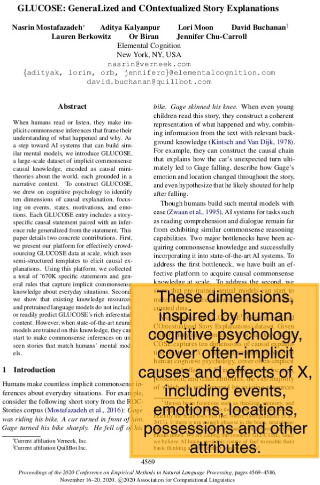
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, emotions, locations, possessions and other attributes.

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)
- captures different dimensions of causal explanation

What is GLUCOSE?



GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh[*] Aditya Kalyanpur Lori Moon David Buchanan[†]
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an inference rule generalized from the statement. This paper details two concrete contributions. First, we present our platform for effectively crowdsourcing GLUCOSE data at scale, which uses semi-structured templates to elicit causal explanations. Using this platform, we collected a total of ~670K specific statements and general rules that capture implicit commonsense knowledge about everyday situations. Second, we show that existing knowledge resources and pretrained language models do not include or readily predict GLUCOSE's rich inferential content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we ...

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

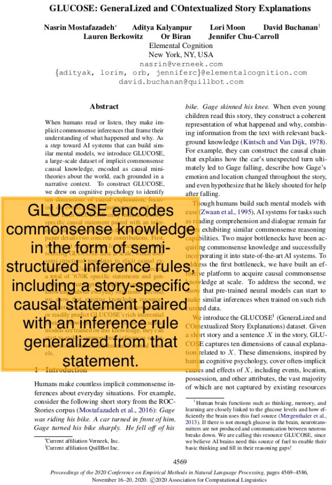
4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586, November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)
- captures different dimensions of causal explanation
- inspired by human cognitive psychology

What is GLUCOSE?



GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh[*] Aditya Kalyanpur Lori Moon David Buchanan[†]
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ...

GLUCOSE encodes commonsense knowledge in the form of semi-structured inference rules, including a story-specific causal statement paired with an inference rule generalized from that statement.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

[*]Current affiliation Verneek, Inc.
[†]Current affiliation QuillBot Inc.

[1]Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

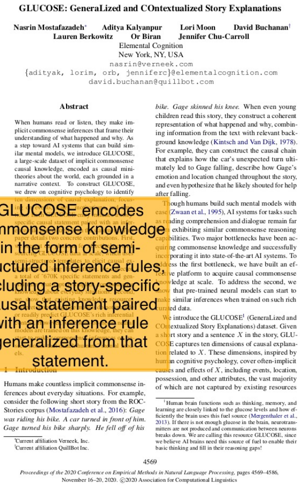
4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)
- captures different dimensions of causal explanation
- inspired by human cognitive psychology

What is GLUCOSE?


GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* Aditya Kalyanpur Lori Moon David Buchanan†
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focus- ... specific causal statement paired with an infer- ... paper details two concrete contributions. First, ... semi-structured templates to elicit causal ex- ... a total of ~670K specific statements and gen- ... or readily predict GLUCOSE's rich inferential ... models are trained on this knowledge, they can ... els.

GLUCOSE encodes commonsense knowledge in the form of semi-structured inference rules, including a story-specific causal statement paired with an inference rule generalized from that statement.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

[1] Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

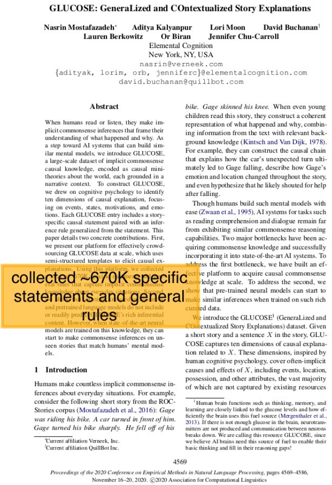
4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)
- captures different dimensions of causal explanation
- inspired by human cognitive psychology
- encodes semi-structured inference rules
  - specific statement
  - generalized rule of statement

What is GLUCOSE?



GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* Aditya Kalyanpur Lori Moon David Buchanan†
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an inference rule generalized from the statement. This paper details two concrete contributions. First, we present our platform for effectively crowdsourcing GLUCOSE data at scale, which uses semi-structured templates to elicit causal explanations. Using this platform, we collected ... and pretrained language models do not include ... content. However, when state-of-the-art neural models are trained on this knowledge, they can start to make commonsense inferences on unseen stories that match humans' mental models.

collected ~670K specific statements and general rules

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

[1] Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

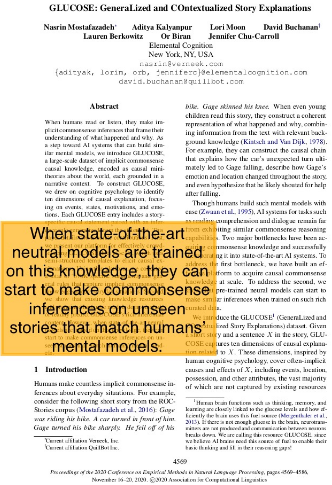
4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586, November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations: (Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)
- captures different dimensions of causal explanation
- inspired by human cognitive psychology
- encodes semi-structured inference rules
  - specific statement
  - generalized rule of statement

What is GLUCOSE?



GLUCOSE: GeneraLized and COntextualized Story Explanations

Nasrin Mostafazadeh* Aditya Kalyanpur Lori Moon David Buchanan†
Lauren Berkowitz Or Biran Jennifer Chu-Carroll
Elemental Cognition
New York, NY, USA
nasrin@verneek.com
{adityak, lorim, orb, jenniferc}@elementalcognition.com
david.buchanan@quillbot.com

Abstract

When humans read or listen, they make implicit commonsense inferences that frame their understanding of what happened and why. As a step toward AI systems that can build similar mental models, we introduce GLUCOSE, a large-scale dataset of implicit commonsense causal knowledge, encoded as causal mini-theories about the world, each grounded in a narrative context. To construct GLUCOSE, we drew on cognitive psychology to identify ten dimensions of causal explanation, focusing on events, states, motivations, and emotions. Each GLUCOSE entry includes a story-specific causal statement paired with an infer- [illegible] we present our platform for effectively crowd- [illegible] semi-structured templates to elicit causal ex- [illegible] eral rules that capture implicit commonsense [illegible] we show that existing knowledge resources [illegible] content. However, when state-of-the-art neural [illegible] start to make commonsense inferences on un- [illegible] els.

1 Introduction

Humans make countless implicit commonsense inferences about everyday situations. For example, consider the following short story from the ROCStories corpus (Mostafazadeh et al., 2016): *Gage was riding his bike. A car turned in front of him. Gage turned his bike sharply. He fell off of his bike. Gage skinned his knee.* When even young children read this story, they construct a coherent representation of what happened and why, combining information from the text with relevant background knowledge (Kintsch and Van Dijk, 1978). For example, they can construct the causal chain that explains how the car's unexpected turn ultimately led to Gage falling, describe how Gage's emotion and location changed throughout the story, and even hypothesize that he likely shouted for help after falling.

Though humans build such mental models with ease (Zwaan et al., 1995), AI systems for tasks such as reading comprehension and dialogue remain far from exhibiting similar commonsense reasoning capabilities. Two major bottlenecks have been acquiring commonsense knowledge and successfully incorporating it into state-of-the-art AI systems. To address the first bottleneck, we have built an effective platform to acquire causal commonsense knowledge at scale. To address the second, we show that pre-trained neural models can start to make similar inferences when trained on such rich curated data.

We introduce the GLUCOSE[1] (GeneraLized and COntextualized Story Explanations) dataset. Given a short story and a sentence X in the story, GLUCOSE captures ten dimensions of causal explanation related to X. These dimensions, inspired by human cognitive psychology, cover often-implicit causes and effects of X, including events, location, possession, and other attributes, the vast majority of which are not captured by existing resources

*Current affiliation Verneek, Inc.
†Current affiliation QuillBot Inc.

[1] Human brain functions such as thinking, memory, and learning are closely linked to the glucose levels and how efficiently the brain uses this fuel source (Mergenthaler et al., 2013). If there is not enough glucose in the brain, neurotransmitters are not produced and communication between neurons breaks down. We are calling this resource GLUCOSE, since we believe AI brains need this source of fuel to enable their basic thinking and fill in their reasoning gaps!

4569

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 4569–4586,
November 16–20, 2020. ©2020 Association for Computational Linguistics

GeneraLized and COntextualized Story Explanations:
(Mostafazadeh et al., 2020)

- large-scale dataset of implicit commonsense causal knowledge
- based on short children's stories from the ROCStories corpus (Mostafazadeh et al., 2016)
- captures different dimensions of causal explanation
- inspired by human cognitive psychology
- encodes semi-structured inference rules
  - specific statement
  - generalized rule of statement
- models trained on this knowledge can start to match humans' mental models

Key Questions

- What kind of knowledge is contained in the GLUCOSE dataset?
- How can we represent that knowledge?
- What power lies in the generalization of that knowledge?
- Is the generalization an factor of improvement?
- How much generalization is too much?

Table of Contents

4584 Annotated Stories

Each story consists of 5 sentences:

Story Example:

1. Tom always loved the ocean.
2. He would vacation at the beach often.
3. He decided to buy a vacation home.
4. He got one right on the beach.
5. Tom spent a lot of time there.

Given a **Sentence** from a **Story**...

Story Example:

1. Tom always loved the ocean.
2. **He would vacation at the beach often.**
3. He decided to buy a vacation home.
4. He got one right on the beach.
5. Tom spent a lot of time there.

...annotate **Inference Rules** in different **Dimensions**:

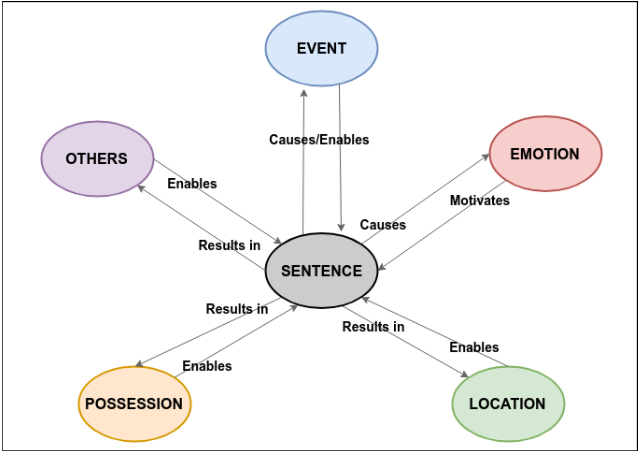


Figure 2: Annotation Dimensions of GLUCOSE Dataset

How does an annotation look like?

Story Example:

1. Tom always loved the ocean.
2. **He would vacation at the beach often.**
3. He decided to buy a vacation home.
4. He got one right on the beach.
5. Tom spent a lot of time there.

Each annotation has two forms:

Specific Rule:

Tom love(s) the ocean `MOTIVATES` Tom vacations at the beach often

General Rule:

Someone_A love(s) the ocean `MOTIVATES` Someone_A vacations at the beach often

How does an annotation look like?

Story Example:

1. Tom always loved the ocean.
2. **He would vacation at the beach often.**
3. He decided to buy a vacation home.
4. He got one right on the beach.
5. Tom spent a lot of time there.

EMOTION

Tom love(s) the ocean	`MOTIVATES`	Tom vacations at the beach often
Someone_A love(s) the ocean	`MOTIVATES`	Someone_A vacations at the beach often

How does an annotation look like?

Story Example:

1. Tom always loved the ocean.
2. **He would vacation at the beach often.**
3. He decided to buy a vacation home.
4. He got one right on the beach.
5. Tom spent a lot of time there.

EMOTION

Tom love(s) the ocean	MOTIVATES	Tom vacations at the beach often
Someone_A love(s) the ocean	MOTIVATES	Someone_A vacations at the beach often

POSSESSION

Tom possess(es) lots of money	ENABLES	Tom vacations at the beach often
Someone_A possess(es) lots of money	ENABLES	Someone_A vacations often

How does an annotation look like?

Story Example:

1. Tom always loved the ocean.
2. **He would vacation at the beach often.**
3. He decided to buy a vacation home.
4. He got one right on the beach.
5. Tom spent a lot of time there.

LOCATION

Tom vacations to the beach	RESULTS IN	Tom is at the beach
Someone_A goes to Somewhere_A	RESULTS IN	Someone_A is at Somewhere_A

How does an annotation look like?

Story Example:

1. Tom always loved the ocean.
2. **He would vacation at the beach often.**
3. He decided to buy a vacation home.
4. He got one right on the beach.
5. Tom spent a lot of time there.

LOCATION

Tom vacations to the beach	`RESULTS IN`	Tom is at the beach
Someone_A goes to Somewhere_A	`RESULTS IN`	Someone_A is at Somewhere_A

EVENT

Tom would vacation at the beach often	`ENABLES`	Tom decided to buy a vacation home
Someone_A would vacation at Somewhere_A often	`ENABLES`	Someone_A decided to buy Something_A

What Knowledge is in the Data?

- Stories form a **Narrative Context**

What Knowledge is in the Data?

- Stories form a **Narrative Context**
- Annotated **Inference Rules**
  - specific
  - general

What Knowledge is in the Data?

- Stories form a **Narrative Context**
- Annotated **Inference Rules**
  - specific
  - general
- **Generalizing Gap Fillers** (something_a, somewhere_a)

What Knowledge is in the Data?

- Stories form a **Narrative Context**
- Annotated **Inference Rules**
  - specific
  - general
- **Generalizing Gap Fillers** (something_a, somewhere_a)
- **Dimension Label**
  - EVENT
  - EMOTION
  - LOCATION
  - POSSESSION
  - OTHERS

What is this Knowledge about?

What is this Knowledge about?

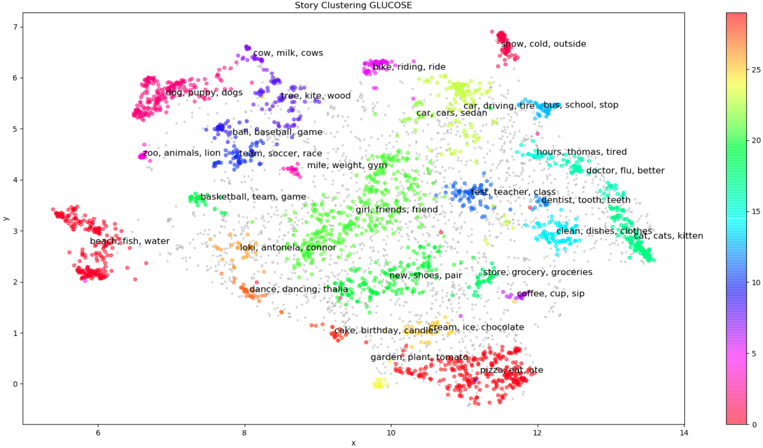


Figure 3: Story Clustering of GLUCOSE with minimal cluster size 15

Table of Contents

Graph Structure for Knowledge Representation

Many **Knowledge Resources** are represented in a **Graph Structure**:

- Google
- Wikipedia
- WordNet, ConceptNet, *Net

Graph Structure for Knowledge Representation

Many **Knowledge Resources** are represented in a **Graph Structure**:

- Google
- Wikipedia
- WordNet, ConceptNet, *Net

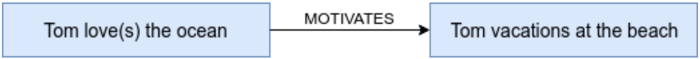
Graphs are the common structure to capture semantic meaning. Concepts (i.e. words, sentences, stories) can be represented by nodes, relations between concepts form edges between nodes.

Graph Structure for Knowledge Representation

Many **Knowledge Resources** are represented in a **Graph Structure**:

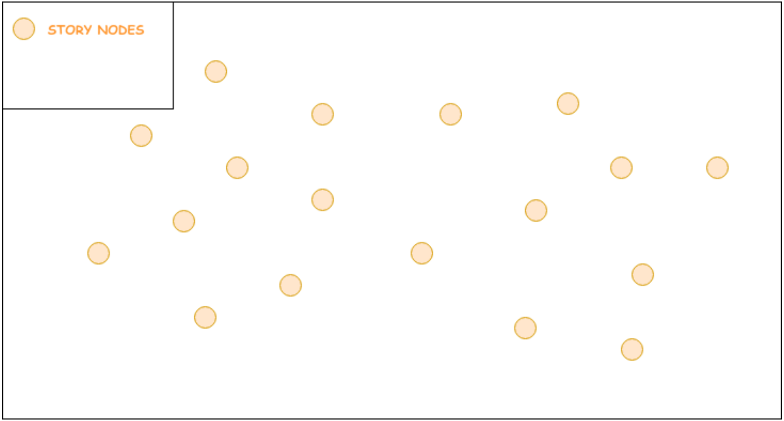
- Google
- Wikipedia
- WordNet, ConceptNet, *Net

Graphs are the common structure to capture semantic meaning. Concepts (i.e. words, sentences, stories) can be represented by nodes, relations between concepts form edges between nodes.



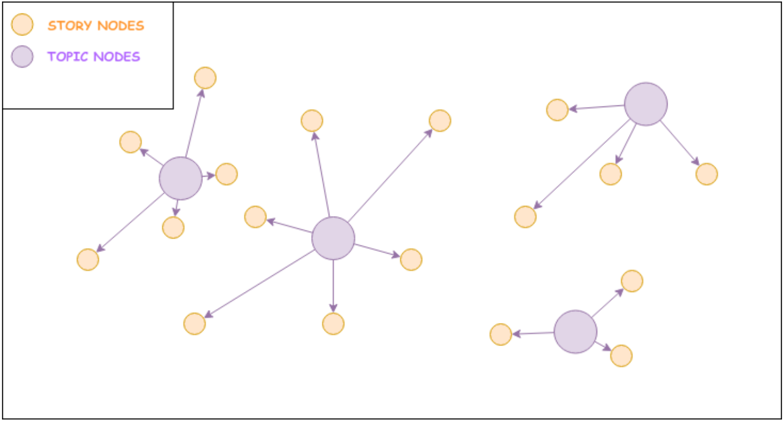
Graph Building Part 1: Knowledge in the Stories

Imagine every story was a node...



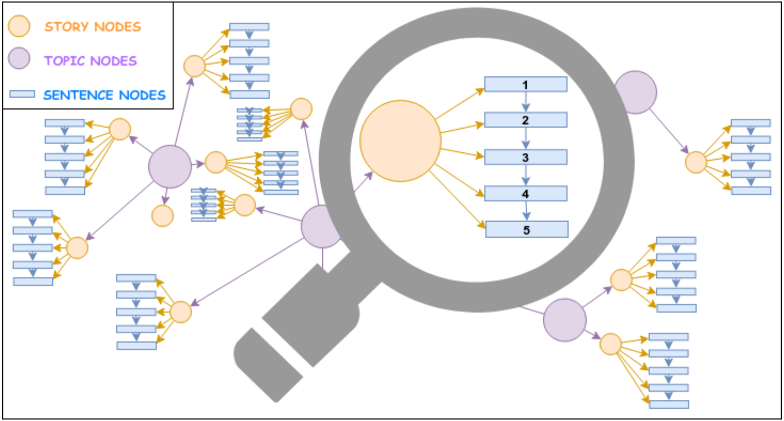
Graph Building Part 1: Knowledge in the Stories

Every STORY NODE is connected to one TOPIC NODE



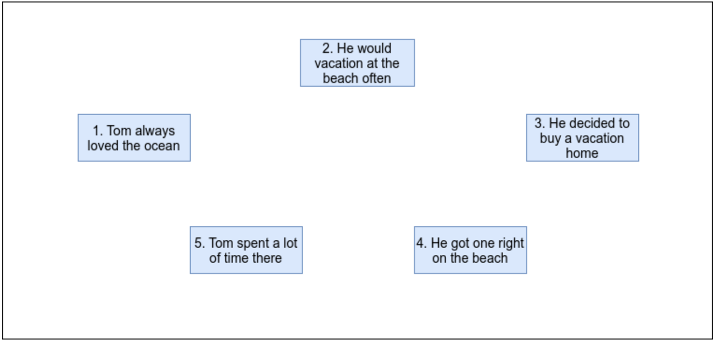
Graph Building Part 1: Knowledge in the Stories

Every STORY NODE points to five SENTENCE NODES



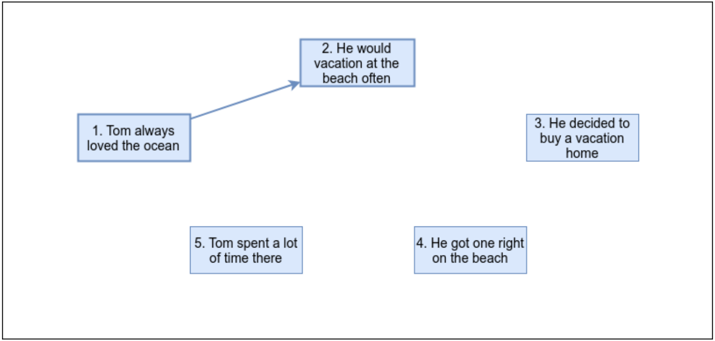
Graph Building Part 2: Knowledge in the Specific Rules

Imagine every sentence was a node...



Graph Building Part 2: Knowledge in the Specific Rules

deriving relations between SENTENCE NODES from specific rules

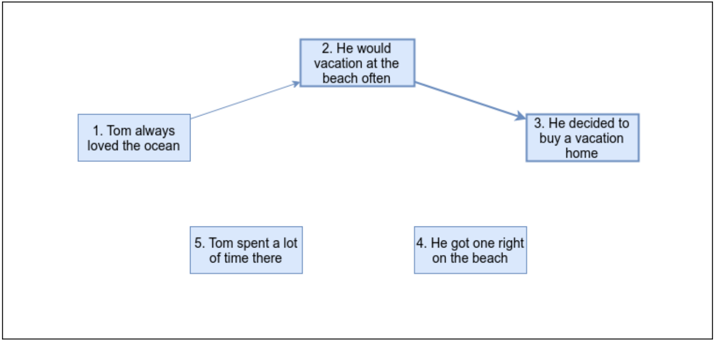


EMOTION

Tom love(s) the ocean `MOTIVATES` Tom vacations at the beach often

Graph Building Part 2: Knowledge in the Specific Rules

deriving relations between SENTENCE NODES from specific rules

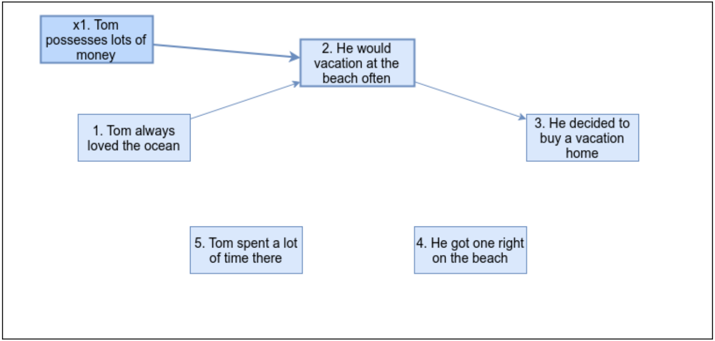


EVENT

Tom would vacation at the beach often `ENABLES` Tom decided to buy a vacation home

Graph Building Part 2: Knowledge in the Specific Rules

deriving relations between SENTENCE NODES from specific rules

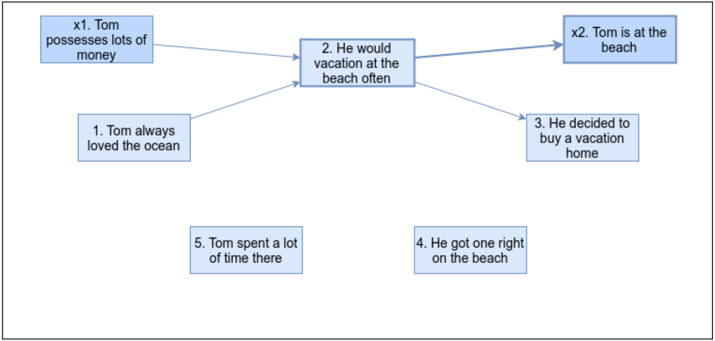


POSSESSION

Tom possess(es) lots of money `ENABLES` Tom vacations at the beach often

Graph Building Part 2: Knowledge in the Specific Rules

deriving relations between SENTENCE NODES from specific rules

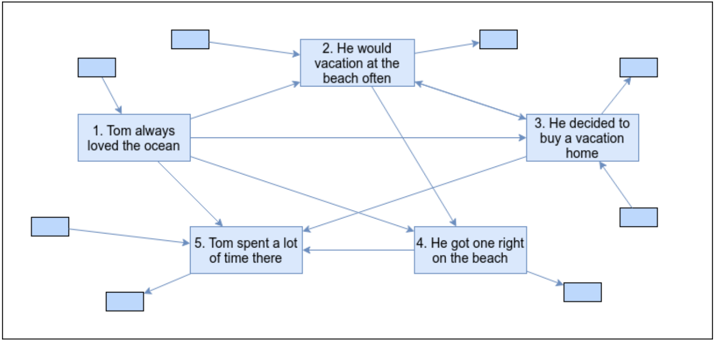


LOCATION

Tom vacations to the beach `RESULTS IN` Tom is at the beach

Graph Building Part 2: Knowledge in the Specific Rules

deriving relations between SENTENCE NODES from specific rules



But what about GENERALIZATION?

Abstract Meaning Representation

Abstract Meaning Representation

- semantic representation language

Abstract Meaning Representation

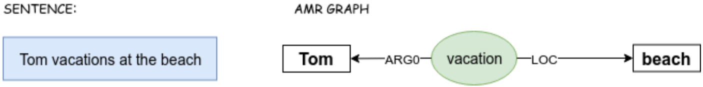
- semantic representation language
- graph structure
  - rooted
  - directed
  - acyclic

Abstract Meaning Representation

- semantic representation language
- graph structure
  - rooted
  - directed
  - acyclic
- include PropBank semantic roles

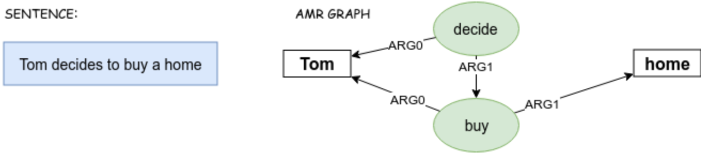
Abstract Meaning Representation

- semantic representation language
- graph structure
  - rooted
  - directed
  - acyclic
- include PropBank semantic roles



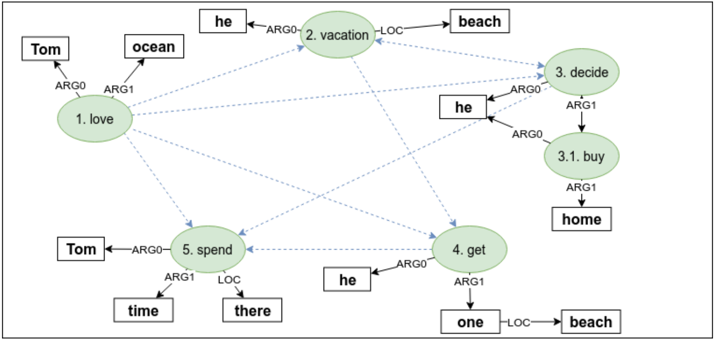
Abstract Meaning Representation

- semantic representation language
- graph structure
  - rooted
  - directed
  - acyclic
- include PropBank semantic roles



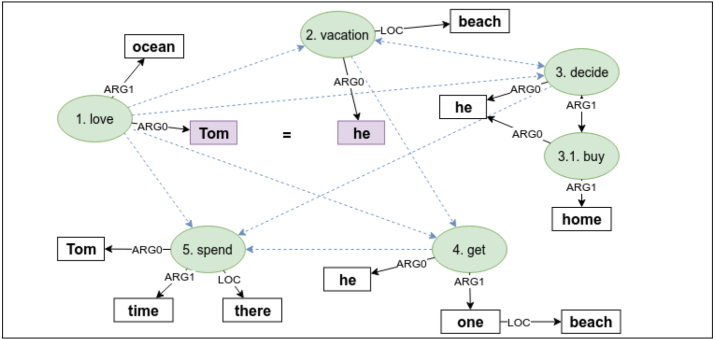
Graph Building Part 3: Knowledge in the General Rules

Imagine every SENTENCE NODE was an AMR GRAPH...



Graph Building Part 3: Knowledge in the General Rules

deriving relations between CONCEPT NODES from general rules

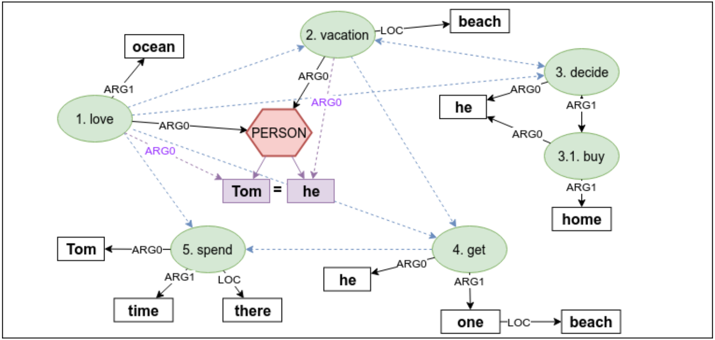


EMOTION

Tom love(s) the ocean	MOTIVATES	Tom vacations at the beach often
Someone_A love(s) the ocean	MOTIVATES	Someone_A vacations at the beach often

Graph Building Part 3: Knowledge in the General Rules

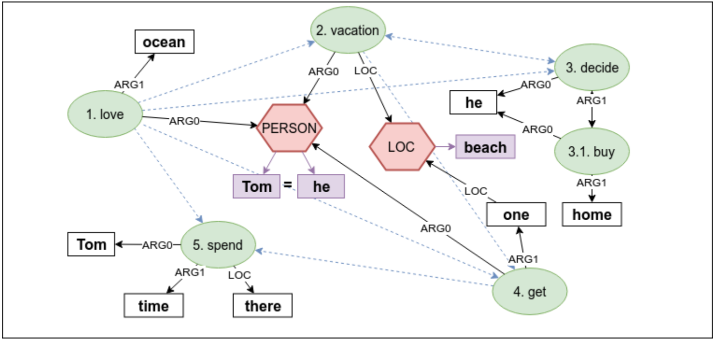
generalizing CONCEPT NODES with Gap Fillers



EMOTION

Tom love(s) the ocean	MOTIVATES	Tom vacations at the beach often
Someone_A love(s) the ocean	MOTIVATES	Someone_A vacations at the beach often

Graph Building Part 3: Knowledge in the General Rules

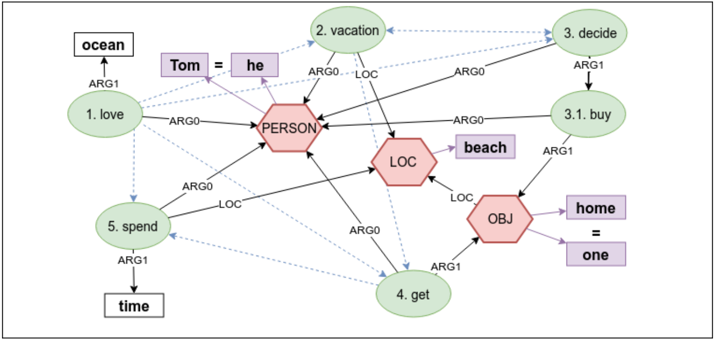


EVENT

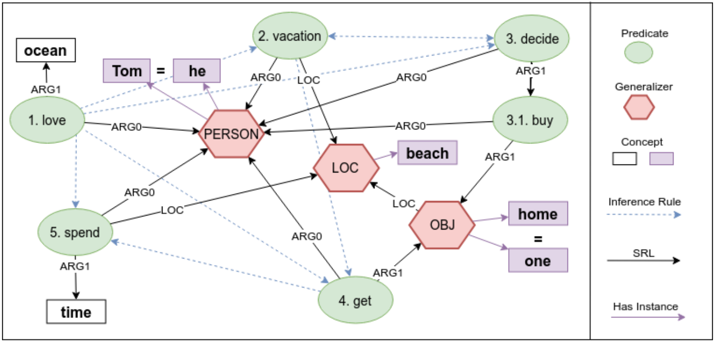
Tom would vacation at the beach often	ENABLES	Tom got a home at the beach
Someone_A would vacation at Somewhere_A often	ENABLES	Someone_A got a home at Somewhere_A

Graph Building Part 3: Knowledge in the General Rules

do that for every rule

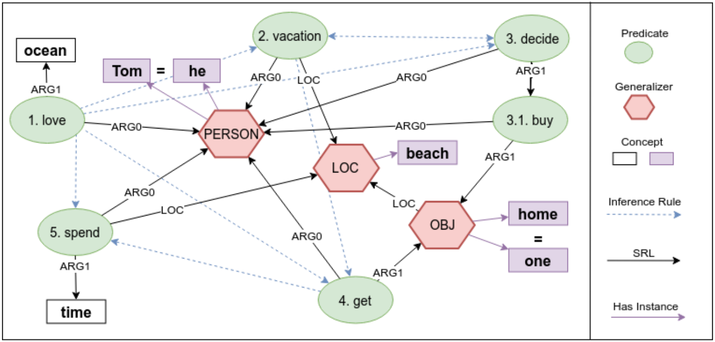


Graph Building Part 3: Knowledge in the General Rules



Graph Building Part 3: Knowledge in the General Rules

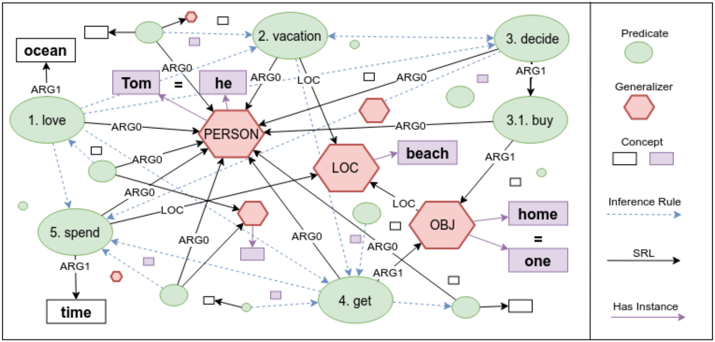
but what about...?



POSSESSION

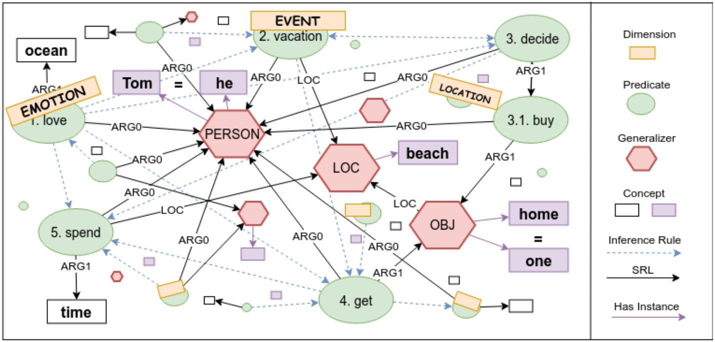
Tom possess(es) lots of money	ENABLES	Tom vacations at the beach often
Someone_A possess(es) lots of money	ENABLES	Someone_A vacations often

Graph Building Part 3: Knowledge in the General Rules

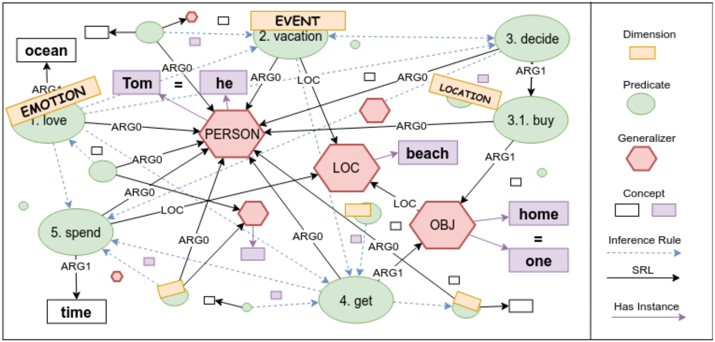


Graph Building Part 3: Knowledge in the General Rules

and don't forget the dimensions label...

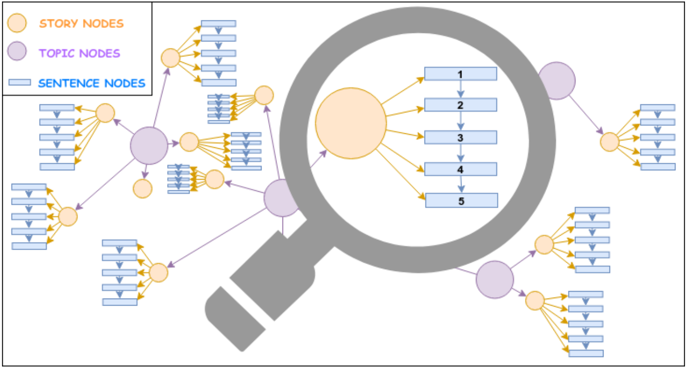


Graph Building Part 3: Knowledge in the General Rules



Now we built one semi-structured graph per story based on annotated semi-structured rules.

Do you remember...?



Do you remember...?

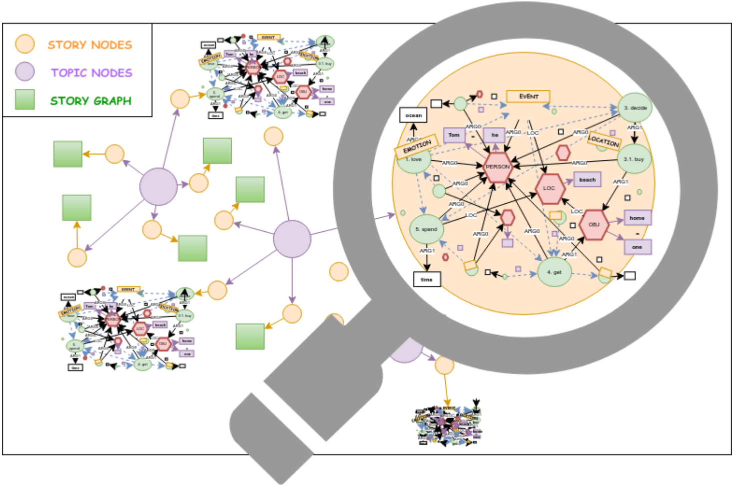
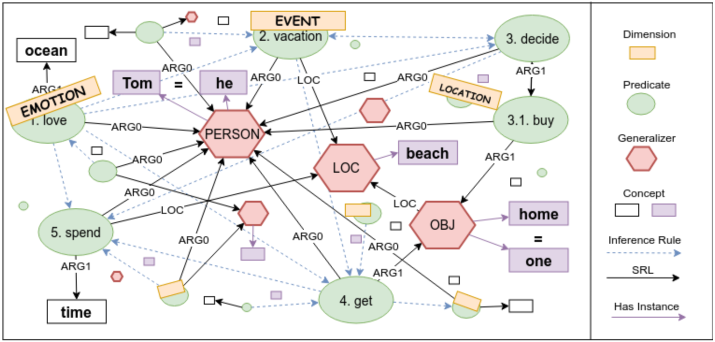


Table of Contents

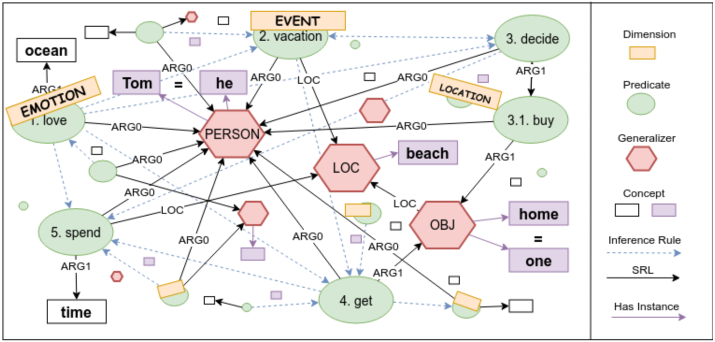
What now?



deduct a...

1. fully instanciated graph
2. highly generalized graph
3. intermediate graph

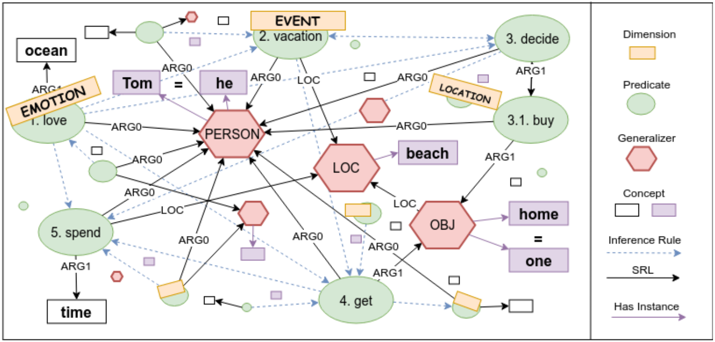
What now?



given these representations, how many times can a rule be applied?

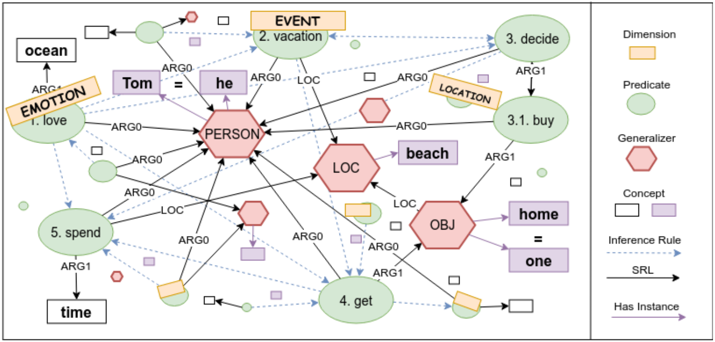
1. Tom love(s) the ocean MOTIVATES Tom vacations at the beach
2. PERS_A love(s) OBJ_A MOTIVATES PERS_A vacations at LOC_A
3. PERS_A love(s) the ocean MOTIVATES PERS_A vacations at the beach

What now?



- how much generalization is a knowledge gain?
- how much generalization is too much?

What now?



- on what and how many concepts can a general rule be applied?
- are there certain patterns depending on topics?
- are there certain patterns depending on dimensions?

THANK YOU FOR YOUR ATTENTION