Low-dimensional Embeddings of Logic

Tim Rocktäschel,[1] Matko Bosnjak,[1] Sameer Singh[2] and Sebastian Riedel[1]

[1] University College London
[2] University of Washington

Machine Reading

ACL 2014 Workshop on Semantic Parsing

26th June 2014

Motivation

Motivation

Machine Reading and Reasoning

Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

- "All mammals are endothermic."

  $\forall x : \textsc{mammal}(x) \Rightarrow \textsc{endothermic}(x)$

Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

- "All mammals are endothermic."

  $\forall x : \textsc{mammal}(x) \Rightarrow \textsc{endothermic}(x)$

- Reasoning...

  $\textsc{endothermic}(\textsc{Colugo})$

Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

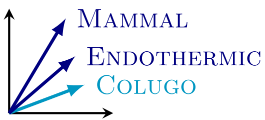
- "All mammals are endothermic."

  $\forall x : \textsc{mammal}(x) \Rightarrow \textsc{endothermic}(x)$

- Reasoning...

  $\textsc{endothermic}(\textsc{Colugo})$

I wish I had a distributed model...



Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

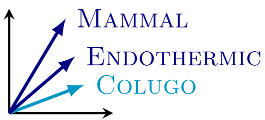
- "All mammals are endothermic."

  $\forall x : \textsc{mammal}(x) \Rightarrow \textsc{endothermic}(x)$

- Reasoning...

  $\textsc{endothermic}(\textsc{Colugo})$

I wish I had a distributed model...



Debugging Distributed Representations

Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

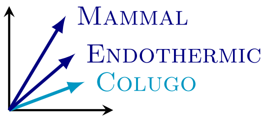
- "All mammals are endothermic."

  $\forall x : \textsc{mammal}(x) \Rightarrow \textsc{endothermic}(x)$

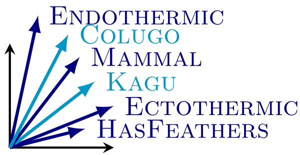
- Reasoning...

  $\textsc{endothermic}(\textsc{Colugo})$

I wish I had a distributed model...



Debugging Distributed Representations



Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

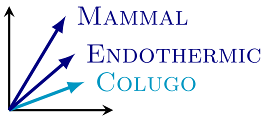
- "All mammals are endothermic."

  $\forall x : \textsc{mammal}(x) \Rightarrow \textsc{endothermic}(x)$

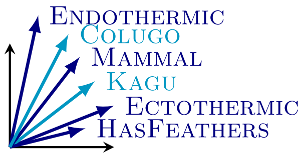
- Reasoning...

  $\textsc{endothermic}(\textsc{Colugo})$

I wish I had a distributed model...



Debugging Distributed Representations



- Wrong Predictions

  $\textsc{mammal}(\textsc{Kagu})$

  $\textsc{ectothermic}(\textsc{Colugo})$

Motivation

Machine Reading and Reasoning

- "Colugos are arboreal gliding mammals that are found in Southeast Asia."

  $\textsc{mammal}(\textsc{Colugo})$

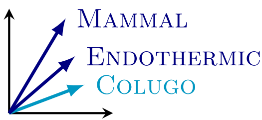
- "All mammals are endothermic."

  $\forall x : \textsc{mammal}(x) \Rightarrow \textsc{endothermic}(x)$

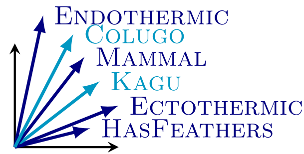
- Reasoning...

  $\textsc{endothermic}(\textsc{Colugo})$

I wish I had a distributed model...



Debugging Distributed Representations



- Wrong Predictions

  $\textsc{mammal}(\textsc{Kagu})$

  $\textsc{ectothermic}(\textsc{Colugo})$

I wish I could fix this with...

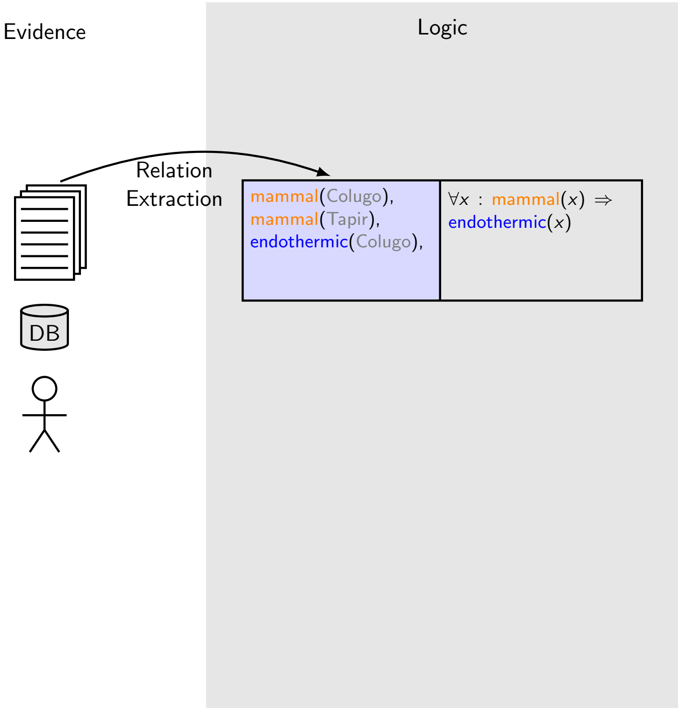
$\forall x : \textsc{hasFeathers}(x) \Rightarrow \neg\textsc{mammal}(x)$

$\forall x : \textsc{animal}(x) \Rightarrow \textsc{endothermic}(x) \oplus \textsc{ectothermic}(x)$

Information Extraction

Evidence

Information Extraction



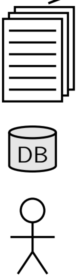
Information Extraction

Evidence

Logic

Discourse Representation Theory or Semantic Parsing

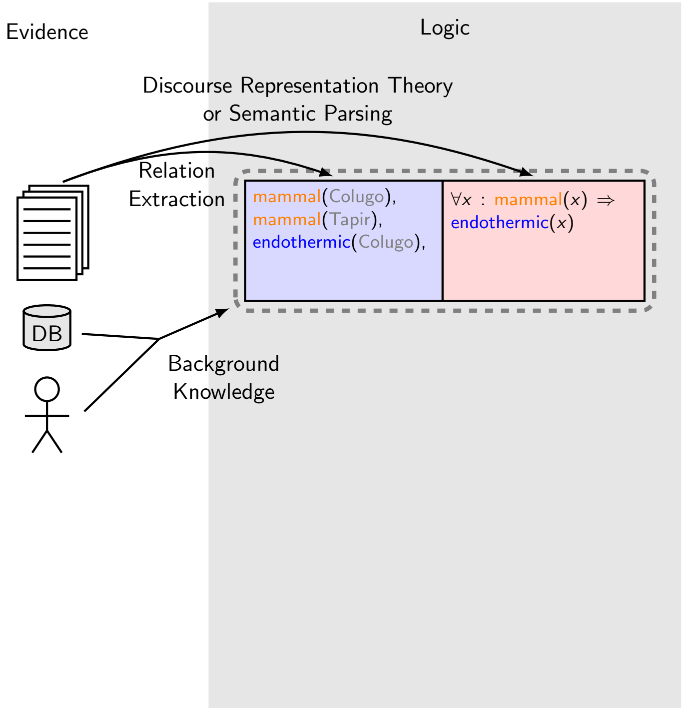
Relation Extraction



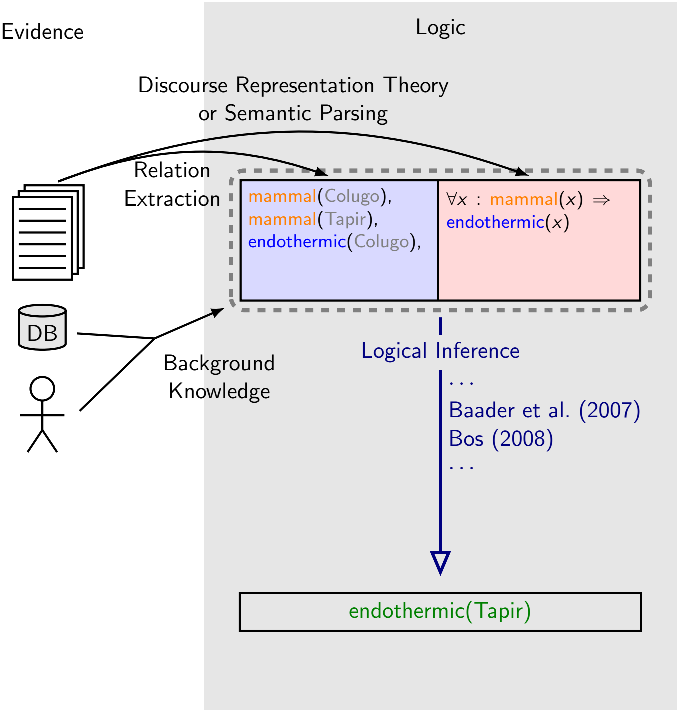
mammal(Colugo),
mammal(Tapir),
endothermic(Colugo),

$\forall x$: mammal(x) $\Rightarrow$ endothermic(x)

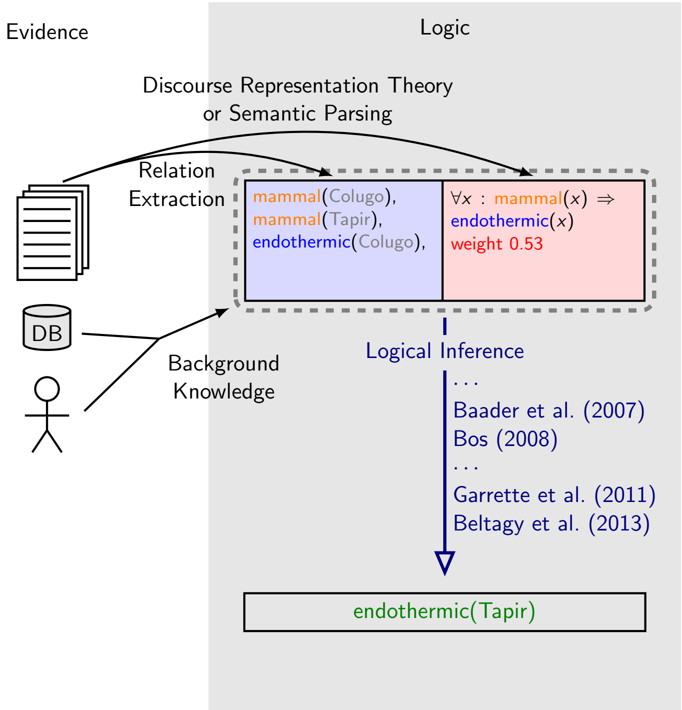
Information Extraction



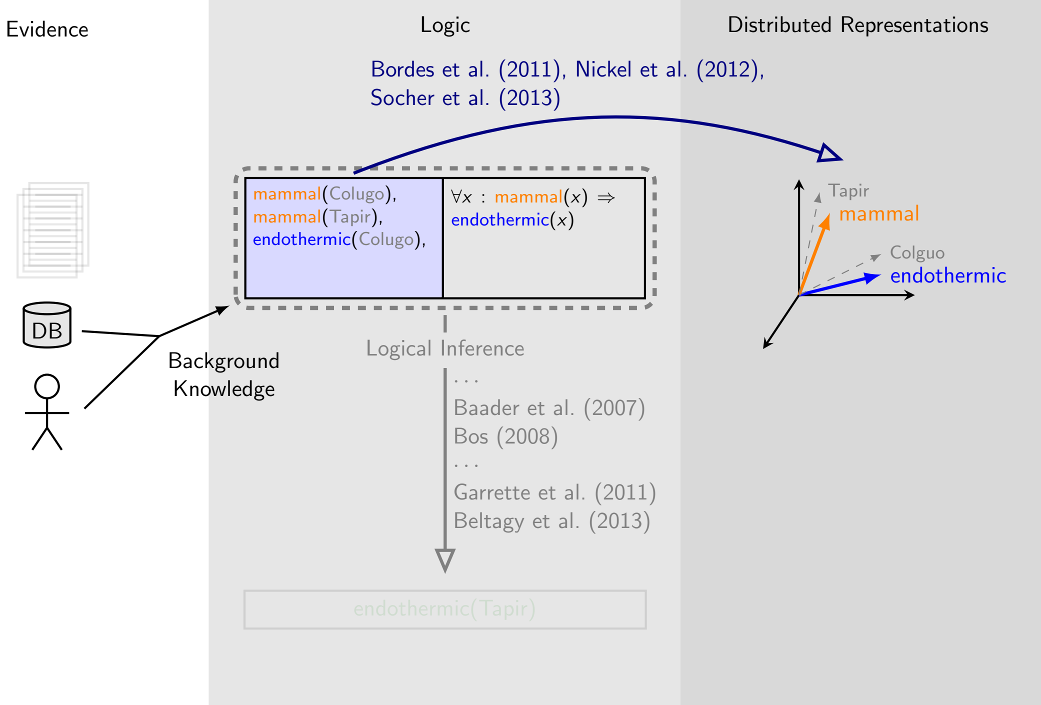
Logical Inference



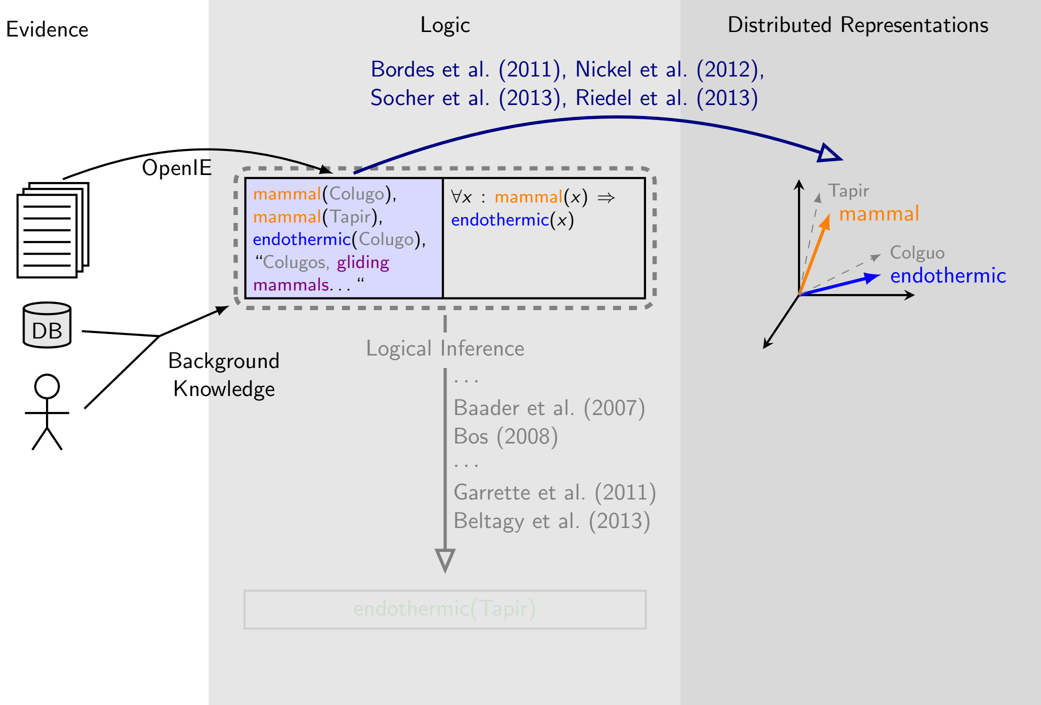
Logical Inference



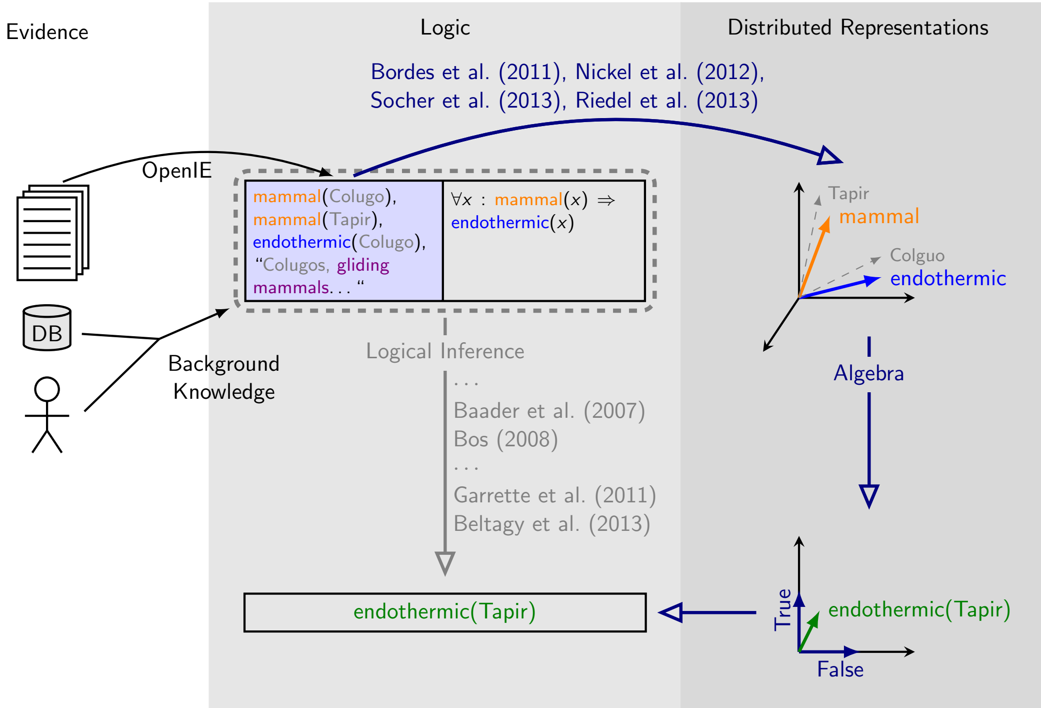
Inference via Distributed Representations



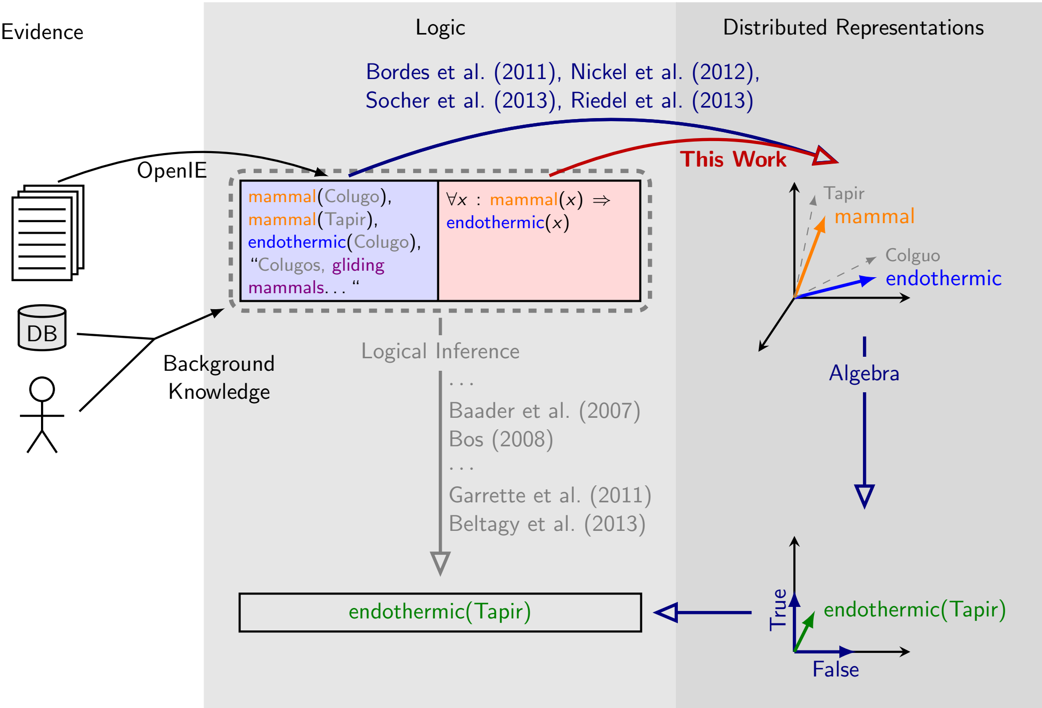
Inference via Distributed Representations



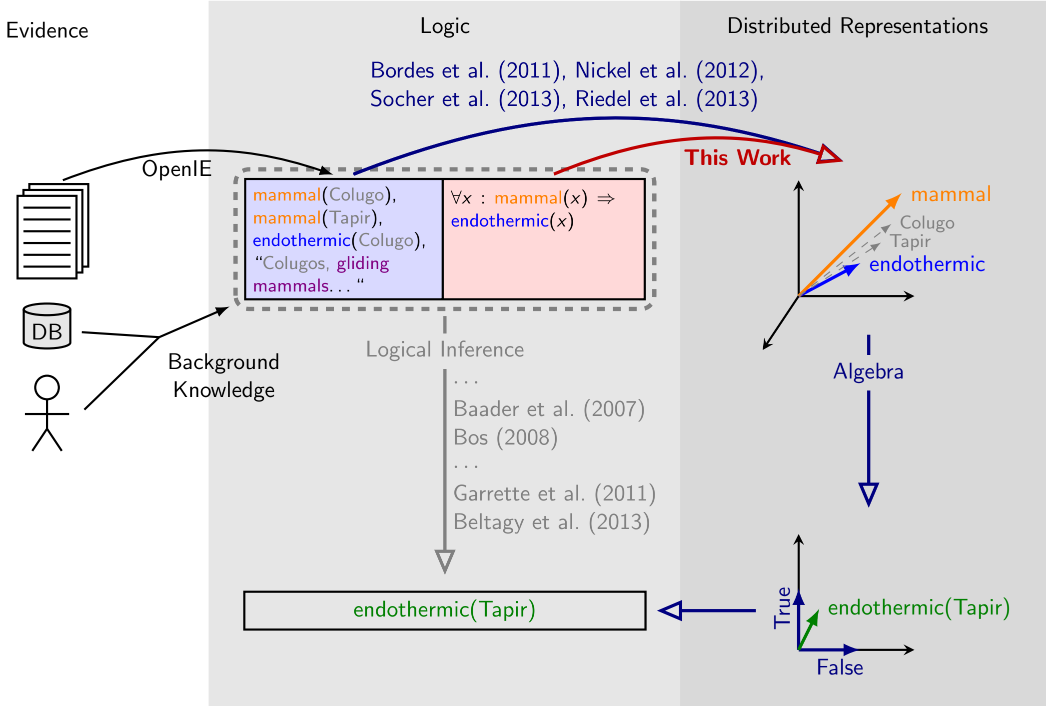
Inference via Distributed Representations



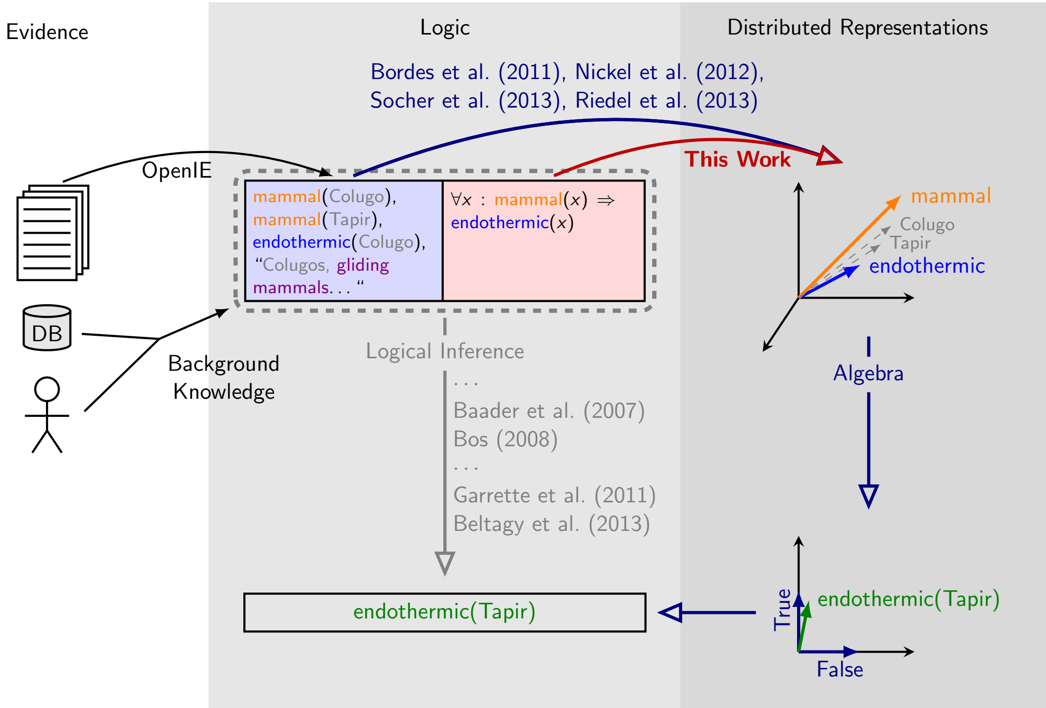
Distributed Representations that Simulate First-order Logical Reasoning



Distributed Representations that Simulate First-order Logical Reasoning



Distributed Representations that Simulate First-order Logical Reasoning



Propositional Logic

Logic	Logical Tensor Calculus (Grefenstette, 2013)
[true]; [false]	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Propositional Logic

Logic	Logical Tensor Calculus (Grefenstette, 2013)
[true]; [false]	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$; $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
[¬]; [∧]; [⇒]	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$; $\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right]$; $\left[\begin{array}{cc\|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$

Propositional Logic

Logic	Logical Tensor Calculus (Grefenstette, 2013)
[true]; [false]	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
[¬]; [∧]; [⇒]	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right]; \left[\begin{array}{cc\|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$
$[\neg\mathcal{A}]$	$[\neg][\mathcal{A}]$

Propositional Logic

Logic	Logical Tensor Calculus (Grefenstette, 2013)
[true]; [false]	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
$[\neg]$; $[\wedge]$; $[\Rightarrow]$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right]; \left[\begin{array}{cc\|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$
$[\neg \mathcal{A}]$	$[\neg][\mathcal{A}]$
$[\mathcal{A} \wedge \mathcal{B}]$	$[\wedge] \times_1 [\mathcal{A}] \times_2 [\mathcal{B}]$
$[\mathcal{A} \Rightarrow \mathcal{B}]$	$[\Rightarrow] \times_1 [\mathcal{A}] \times_2 [\mathcal{B}]$

Propositional Logic

Logic	Logical Tensor Calculus (Grefenstette, 2013)
[true]; [false]	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
[¬]; [∧]; [⇒]	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right]; \left[\begin{array}{cc\|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$
$[\neg\mathcal{A}]$	$[\neg][\mathcal{A}]$
$[\mathcal{A} \wedge \mathcal{B}]$	$[\wedge] \times_1 [\mathcal{A}] \times_2 [\mathcal{B}]$
$[\mathcal{A} \Rightarrow \mathcal{B}]$	$[\Rightarrow] \times_1 [\mathcal{A}] \times_2 [\mathcal{B}]$
$[\mathcal{A} \wedge \neg\mathcal{B} \Rightarrow \neg\mathcal{C}]$	$[\Rightarrow] \times_1 ([\wedge] \times_1 [\mathcal{A}] \times_2 [\neg][\mathcal{B}]) \times_2 [\neg][\mathcal{C}]$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)
[Colugo]	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^\mathsf{T}$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)
[$\textsc{Colugo}$]	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^{\mathsf{T}}$
[$\textsc{mammal}$]	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)
[COLUGO]	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^{\mathsf{T}}$
[MAMMAL]	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$
[MAMMAL(COLUGO)]	[MAMMAL][COLUGO]

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^T$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X \rvert = \\ \quad \lvert \{x \mid F(x) = [\text{true}]\} \rvert \\ [\text{false}] \text{ otherwise} \end{cases}$
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if} \\ \quad \lvert \{x \mid F(x) = [\text{true}]\} \rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^\mathsf{T}$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X\rvert = \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert \\ [\text{false}] \text{ otherwise} \end{cases}$
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if} \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$

Example: $Q = \forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)	Low-Rank Distributed Representation
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^\mathsf{T}$	$\begin{bmatrix} -0.05 & 0.44 & 1.38 \end{bmatrix}^\mathsf{T}$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -0.56 & 0.24 & 0.63 \\ 0.12 & 0.93 & -0.16 \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$	
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X\rvert = \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert \\ [\text{false}] \text{ otherwise} \end{cases}$	
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if} \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$	

Example: $Q = \forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)	Low-Rank Distributed Representation
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^\mathsf{T}$	$\begin{bmatrix} -0.05 & 0.44 & 1.38 \end{bmatrix}^\mathsf{T}$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -0.56 & 0.24 & 0.63 \\ 0.12 & 0.93 & -0.16 \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$	$[\textsc{mammal}][\textsc{Colugo}]$
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X\rvert = \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert \\ [\text{false}] \text{ otherwise} \end{cases}$	
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$	

Example: $Q = \forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)	Low-Rank Distributed Representation
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^\mathsf{T}$	$\begin{bmatrix} -0.05 & 0.44 & 1.38 \end{bmatrix}^\mathsf{T}$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -0.56 & 0.24 & 0.63 \\ 0.12 & 0.93 & -0.16 \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$	$[\textsc{mammal}][\textsc{Colugo}]$
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X\rvert = \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert \\ [\text{false}] \text{ otherwise} \end{cases}$	$\frac{1}{\lvert X\rvert}\sum_x [F(x)]$
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$	

Example: $Q = \forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)	Low-Rank Distributed Representation
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^\mathsf{T}$	$\begin{bmatrix} -0.05 & 0.44 & 1.38 \end{bmatrix}^\mathsf{T}$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -0.56 & 0.24 & 0.63 \\ 0.12 & 0.93 & -0.16 \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$	$[\textsc{mammal}][\textsc{Colugo}]$
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X\rvert = \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert \\ [\text{false}] \text{ otherwise} \end{cases}$	$\frac{1}{\lvert X\rvert} \sum_x [F(x)]$
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if} \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$	$[\neg\forall x \in X : \neg F(x)]$

Example: $Q = \forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)	Low-Rank Distributed Representation
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^T$	$\begin{bmatrix} -0.05 & 0.44 & 1.38 \end{bmatrix}^T$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -0.56 & 0.24 & 0.63 \\ 0.12 & 0.93 & -0.16 \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$	$[\textsc{mammal}][\textsc{Colugo}]$
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X\rvert = \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert \\ [\text{false}] \text{ otherwise} \end{cases}$	$\frac{1}{\lvert X\rvert} \sum_x [F(x)]$
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if} \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$	$[\neg \forall x \in X : \neg F(x)]$

Example: $Q = \forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$
confidence$(Q) := [Q]_{(1)}$

Constants, Predicates and Quantifiers

	Full-Rank One-Hot Representation (Grefenstette, 2013)	Low-Rank Distributed Representation
$[\textsc{Colugo}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^\top$	$\begin{bmatrix} ? & ? & ? \end{bmatrix}^\top$
$[\textsc{mammal}]$	$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \end{bmatrix}$
$[\textsc{mammal}(\textsc{Colugo})]$	$[\textsc{mammal}][\textsc{Colugo}]$	$[\textsc{mammal}][\textsc{Colugo}]$
$[\forall x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \lvert X\rvert = \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert \\ [\text{false}] \text{ otherwise} \end{cases}$	$\frac{1}{\lvert X\rvert} \sum_x [F(x)]$
$[\exists x \in X : F(x)]$	$\begin{cases} [\text{true}] \text{ if } \\ \quad \lvert\{x \mid F(x) = [\text{true}]\}\rvert > 0 \\ [\text{false}] \text{ otherwise} \end{cases}$	$[\neg\forall x \in X : \neg F(x)]$

Example: $Q = \forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$
confidence$(Q) := [Q]_{(1)}$

Objective

Find embeddings for predicates and constants such that true (factual or first-order) formulae evaluate to [true] in the vector space.

Objective

Find embeddings for predicates and constants such that true (factual or first-order) formulae evaluate to [true] in the vector space.

- $\mathcal{E}$: Set of entity (or entity-pair) vectors
- $\mathcal{R}$: Set of relation matrices
- $\mathfrak{K}$: Set of logical formulae Q with training signal $\gamma \in \{[\text{true}], [\text{false}]\}$
  - In previous work this only contained factual statements
  - In addition our objective includes first-order logic formulae!
- $\mathcal{L}$: Loss function, e.g., $\|[Q] - \gamma\|_2$

Objective

Find embeddings for predicates and constants such that true (factual or first-order) formulae evaluate to [true] in the vector space.

$\mathcal{E}$: Set of entity (or entity-pair) vectors

$\mathcal{R}$: Set of relation matrices

$\mathfrak{K}$: Set of logical formulae Q with training signal $\gamma \in \{[\text{true}], [\text{false}]\}$

- In previous work this only contained factual statements
- In addition our objective includes first-order logic formulae!

$\mathcal{L}$: Loss function, e.g., $\|[Q] - \gamma\|_2$

$$\min_{[e]\in\mathcal{E},[r]\in\mathcal{R}} \sum_{(Q,\gamma)\in\mathfrak{K}} \mathcal{L}\left([Q], \gamma\right)$$

Objective

Find embeddings for predicates and constants such that true (factual or first-order) formulae evaluate to [true] in the vector space.

$\mathcal{E}$: Set of entity (or entity-pair) vectors

$\mathcal{R}$: Set of relation matrices

$\mathfrak{K}$: Set of logical formulae Q with training signal $\gamma \in \{[\text{true}], [\text{false}]\}$

- In previous work this only contained factual statements
- In addition our objective includes first-order logic formulae!

$\mathcal{L}$: Loss function, e.g., $\|[Q] - \gamma\|_2$

$$\min_{[e]\in\mathcal{E},[r]\in\mathcal{R}} \sum_{(Q,\gamma)\in\mathfrak{K}} \mathcal{L}\left([Q], \gamma\right)$$

Gradients: Backpropagation through expression tree of $\mathcal{F}$

Learning: Stochastic Gradient Descent

Example

	MAMMAL	ENDOTHERMIC	VERTERBRATE
CHIMPANZEE	**1.0**	**1.0**	**1.0**
KOALA	**1.0**	**1.0**	**1.0**
COLUGO	**1.0**	?	?
KAGU	?	**1.0**	?
DODO	?	**1.0**	?

Example

	MAMMAL	ENDOTHERMIC	VERTERBRATE
CHIMPANZEE	**1.0**	**1.0**	**1.0**
KOALA	**1.0**	**1.0**	**1.0**
COLUGO	**1.0**	0.1	0.5
KAGU	0.1	**0.9**	0.3
DODO	0.1	**1.0**	0.3

Example

$\forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

	MAMMAL	ENDOTHERMIC	VERTERBRATE
CHIMPANZEE	**1.0**	**1.0**	**1.0**
KOALA	**1.0**	**1.0**	**1.0**
COLUGO	**1.0**	0.1	0.5
KAGU	0.1	**0.9**	0.3
DODO	0.1	**1.0**	0.3

Example

$\forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

	MAMMAL	ENDOTHERMIC	VERTERBRATE
CHIMPANZEE	**1.0**	**0.9**	**1.0**
KOALA	**1.0**	**0.9**	**1.0**
COLUGO	**0.9**	0.6	0.8
KAGU	0.1	**1.0**	0.2
DODO	0.0	**1.0**	0.1

Example

$\forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

	MAMMAL	ENDOTHERMIC	VERTERBRATE
CHIMPANZEE	**1.0**	**0.9**	**1.0**
KOALA	**1.0**	**0.9**	**1.0**
COLUGO	**0.9**	(0.1) 0.6	(0.5) 0.8
KAGU	0.1	**1.0**	0.2
DODO	0.0	**1.0**	0.1

Example

$\forall x \in X : [\Rightarrow] \times_1 ([\textsc{mammal}][x]) \times_2 ([\textsc{endothermic}][x])$

	MAMMAL	ENDOTHERMIC	VERTERBRATE
Chimpanzee	**1.0**	**0.9**	**1.0**
Koala	**1.0**	**0.9**	**1.0**
Colugo	**0.9**	(0.1) 0.6	(0.5) 0.8
Kagu	0.1	**1.0**	0.2
Dodo	0.0	**1.0**	0.1

Conclusion and Open Questions

✓ Inject first-order logic into vector spaces

Conclusion and Open Questions

- ✓ Inject first-order logic into vector spaces
- ✓ Objective that encourages distributed representations to simulate first-order logical reasoning

Conclusion and Open Questions

- ✓ Inject first-order logic into vector spaces
- ✓ Objective that encourages distributed representations to simulate first-order logical reasoning

- **?** What are the theoretical limits of embedding logical formulae in vector spaces?

Conclusion and Open Questions

✓ Inject first-order logic into vector spaces

✓ Objective that encourages distributed representations to simulate first-order logical reasoning

? What are the theoretical limits of embedding logical formulae in vector spaces?

? What are efficient ways of injecting quantified formulae without iterating over all elements of a domain?

Conclusion and Open Questions

- ✓ Inject first-order logic into vector spaces
- ✓ Objective that encourages distributed representations to simulate first-order logical reasoning

- ? What are the theoretical limits of embedding logical formulae in vector spaces?
- ? What are efficient ways of injecting quantified formulae without iterating over all elements of a domain?
- ? Can we provide provenance of proofs of answers?

Thank you!

References

Franz Baader, Bernhard Ganter, Baris Sertkaya, and Ulrike Sattler. 2007. Completing description logic knowledge bases using formal concept analysis. In IJCAI, pages 230–235.

Islam Beltagy, Cuong Chau, Gemma Boleda, Dan Garrette, Katrin Erk, and Raymond Mooney. 2013. Montague meets markov: Deep semantics with probabilistic logical form. In 2nd Joint Conference on Lexical and Computational Semantics: Proceeding of the Main Conference and the Shared Task, Atlanta, pages 11–21.

Antoine Bordes, Jason Weston, Ronan Collobert, and Yoshua Bengio. 2011. Learning structured embeddings of knowledge bases. In AAAI.

Johan Bos. 2008. Wide-coverage semantic analysis with boxer. In Johan Bos and Rodolfo Delmonte, editors, Semantics in Text Processing. STEP 2008 Conference Proceedings, Research in Computational Semantics, pages 277–286. College Publications.

Dan Garrette, Katrin Erk, and Raymond Mooney. 2011. Integrating logical representations with probabilistic information using markov logic. In Proceedings of the Ninth International Conference on Computational Semantics, pages 105–114. Association for Computational Linguistics.

Edward Grefenstette. 2013. Towards a formal distributional semantics: Simulating logical calculi with tensors. In Proceedings of the Second Joint Conference on Lexical and Computational Semantics, pages 1–10.

Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. 2012. Factorizing yago: scalable machine learning for linked data. In Proc. of WWW, pages 271–280.

Sebastian Riedel, Limin Yao, Andrew McCallum, and Benjamin M Marlin. 2013. Relation extraction with matrix factorization and universal schemas. In Proceedings of NAACL-HLT, pages 74–84.

Richard Socher, Danqi Chen, Christopher D. Manning, and Andrew Y. Ng. 2013. Reasoning with neural tensor networks for knowledge base completion. In NIPS, pages 926–934.