UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR INFORMATIONSSYSTEME

Özgür L. Özçep

Knowledge Graph Embeddings

KI Kolloqium
December 2, 2019

Motivation

- Past talk of Ralf: tour from numbers to logic
- Today: tour from logic to numbers
  - Logic as a useful method/tool for machine learning (ML)

Motivation

- Past talk of Ralf: tour from numbers to logic
- Today: tour from logic to numbers
  - Logic as a useful method/tool for machine learning (ML)
  - Maybe it is more: foundations of ML
  - But not easy ...*)

*)The Burden of Logical Foundations

For all that, PM is not widely used today: probably the foremost reason for this is its reputation for typographical complexity. Somewhat infamously, several hundred pages of PM precede the proof of the validity of the proposition 1+1=2.

Wikipedia entry for Principia Mathematica (accessed November 25, 2019)

Expressivity of Embeddings

Knowledge Graph (KG)

- Set of assertions in triple form $KG = \{ \ (subj \ pred \ obj) \ \}$
- Convenient logical notation $KG = \{ \ pred(subj, obj) \ \}$

Knowledge Graph (KG)

- Set of assertions in triple form $KG = \{ \ (subj \ pred \ obj) \ \}$
- Convenient logical notation $KG = \{ \ pred(subj, obj) \ \}$

Example

$$KG_1 = \{\textit{worksIn(alice,AI)}, \textit{subfield(AI,CS)}, \textit{manyPubls(alice,bob)}\}$$

Knowledge Graph (KG)

- Set of assertions in triple form $KG = \{\ (subj\ pred\ obj)\ \}$
- Convenient logical notation $KG = \{\ pred(subj, obj)\ \}$

Example

$$KG_1 = \{\mathit{worksIn(alice,AI)}, \mathit{subfield(AI,CS)}, \mathit{manyPubls(alice,bob)}\}$$

- Using triples as "atoms" of knowledge

Knowledge Graph (KG)

- Set of assertions in triple form $KG = \{ \ (subj \ pred \ obj) \ \}$
- Convenient logical notation $KG = \{ \ pred(subj, obj) \ \}$

Example

$$KG_1 = \{worksIn(alice,AI), subfield(AI,CS), manyPubls(alice,bob)\}$$

- Using triples as "atoms" of knowledge
- RDF triples

Knowledge Graph (KG)

- Set of assertions in triple form $KG = \{\ (subj\ pred\ obj)\ \}$
- Convenient logical notation $KG = \{\ pred(subj, obj)\ \}$

Example

$$KG_1 = \{worksIn(alice, AI), subfield(AI, CS), manyPubls(alice, bob)\}$$

- Using triples as "atoms" of knowledge
- RDF triples

  W3C RDF

- Analytical philosophy
  Bernard Bolzano (1781-1848)

Knowledge Graph (KG)

Usually KGs highly incomplete $\Longrightarrow$ "Learn" new triples assuming regularities

Example

$$KG_2 = \{\mathit{worksIn(alice,AI)}, \mathit{subfield(AI,CS)}, \mathit{manyPubls(alice,bob)} \\ \underline{\mathit{worksIn(bob,AI)}}\}\}$$

Knowledge Graph Embeddings

- Capture regularities by embedding KG into continuous space $E = \mathbb{R}^n$
- Various recent approaches
  - TRansE, TransR, STransH
  - DistMult
  - ComplEx
  - SimplE
  - RESCAL ...

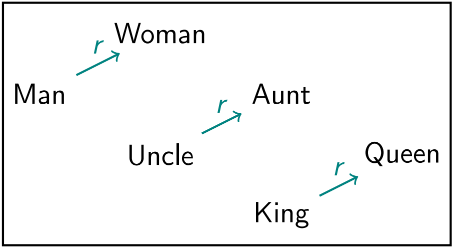
Knowledge Graph Embeddings

- Capture regularities by embedding KG into continuous space $E = \mathbb{R}^n$
- Various recent approaches
  - TRansE, TransR, STransH
  - DistMult
  - ComplEx
  - SimplE
  - RESCAL ...
- Can be described uniformly by scoring functions for relations R

$$s_R : E \times E \longrightarrow \mathbb{R}$$

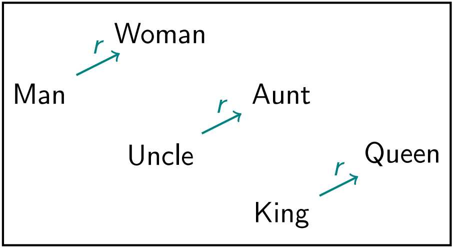
Convention:
$s_R(u, v)$ small means probability of $R(u, v)$ is high.

TransE (Bordes et al. 13)



- $s_R(u, v) = \|u + r - v\|$ ($\|\cdot\|$: Euclidean Norm)

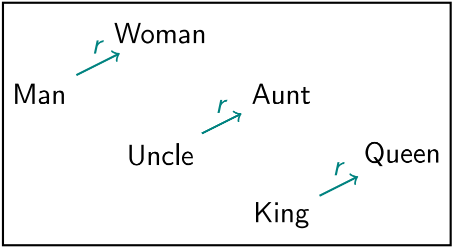
TransE (Bordes et al. 13)



- $s_R(u, v) = \|u + r - v\|$ ($\|\cdot\|$: Euclidean Norm)

- Limitation: Relations r = vector translations, hence functional

TransE (Bordes et al. 13)



- $s_R(u, v) = \|u + r - v\|$ ($\|\cdot\|$: Euclidean Norm)
- Limitation: Relations r = vector translations, hence functional
- Generalizations
  - TransR (map in R-specific space before translation)

    $$s_R(u, v) = \|M_r u + r - M_r v\|$$

  - STransE (different matrices for subject (s) and object (o))

    $$s_R(u, v) = \|M_r^s u + r - M_r^o v\|$$

DistMult (Yang et al. 14)

- $s_R(u, v) = -\sum_{i=1}^{n} u_i r_i v_i$

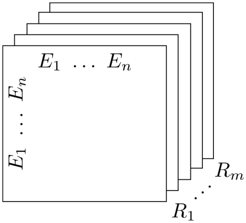
DistMult (Yang et al. 14)

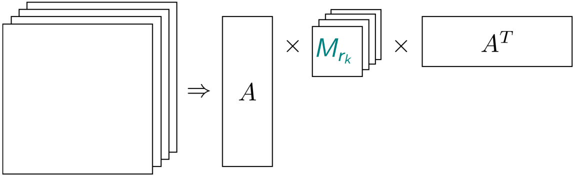
- $s_R(u, v) = -\sum_{i=1}^{n} u_i r_i v_i$
- Limitation: can model only symmetric relations

DistMult (Yang et al. 14)

- $s_R(u, v) = -\sum_{i=1}^{n} u_i r_i v_i$
- Limitation: can model only symmetric relations
- Generalizations
  - Distinguish whether entities in subject or object position
  - ComplEx (Trouillon et al. 16)
    relate those entity types by complex conjugation
  - SimplE (Kazemi/Poole 18)
    for each $R(u, v)$ consider $R^{-}(v, u)$

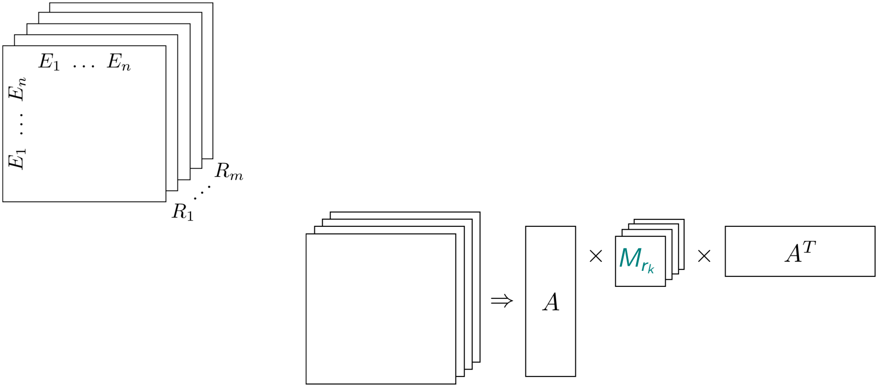
RESCAL





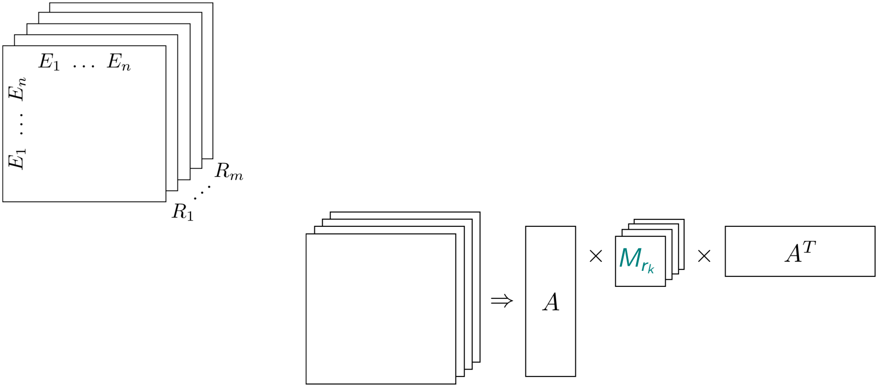
- $s_{R_k}(u, v) = -u^T M_{r_k} v$

RESCAL



- $s_{R_k}(u, v) = -u^T M_{r_k} v$

RESCAL



- $s_{R_k}(u, v) = -u^T M_{r_k} v$
- Limitation: suffers from overfitting

Expressivity Criterion

- Usually one considers low-dimensional spaces
- But nonetheless must be sufficiently high to embed knowledge expressed in KGs
  - P: set of valid triples in KG
  - N: set of non-valid triples in KG $(N \cap P = \emptyset)$

Expressivity Criterion

- Usually one considers low-dimensional spaces
- But nonetheless must be sufficiently high to embed knowledge expressed in KGs
  - P: set of valid triples in KG
  - N: set of non-valid triples in KG $(N \cap P = \emptyset)$

Definition (Kazemi/Poole 18)

An embedding model is fully expressive iff there are a dimension n, an embedding e, and a threshold λ_R such that:

- for all $R(u, v) \in P$: $s_R(e(u), e(v)) \leq \lambda_R$
- for all $R(u, v) \in N$: $s_R(e(u), e(v)) > \lambda_R$

Results on Expressiveness (Kazemi/Poole 18)

Model	Fully expressive?
TransE (and all extensions)	-
DistMult	-
SimplE	+
ComplEx	+
RESCAL	+

- TransE not sufficiently general, RESCAL tends to overfitting
- And now?

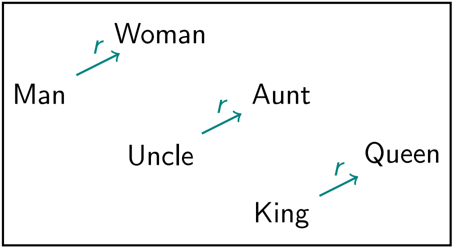
Results on Expressiveness (Kazemi/Poole 18)

Model	Fully expressive?
TransE (and all extensions)	-
DistMult	-
SimplE	+
ComplEx	+
RESCAL	+

- TransE not sufficiently general, RESCAL tends to overfitting
- And now?
- Here comes logic
  - "Logic focussed" semantics for concepts and relations
  - Add logically specified background knowledge

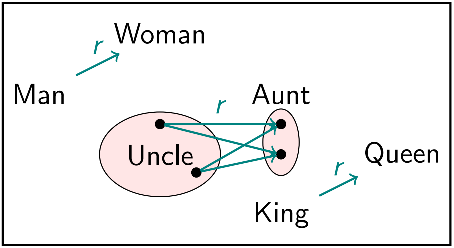
Logico-geometrical Semantics

- Represent concepts as sets (set of vectors, not single vector)
- Represent binary relations as sets of pairs of objects



Logico-geometrical Semantics

- Represent concepts as sets (set of vectors, not single vector)
- Represent binary relations as sets of pairs of objects



Adding Background Knowledge

Some completions due to background knowledge/ontology

Example

Ontology :

$$\{\forall X, Y, Z. subfield(X, Y) \wedge worksIn(Z, X) \rightarrow worksIn(Z, Y)\}$$

$$\begin{aligned} KG_2 \quad = \quad & \{worksIn(alice,AI), subfield(AI,CS), manyPubls(alice,bob), \\ & worksIn(bob,AI), \qquad \text{(by induction)} \\ & worksIn(alice,CS) \qquad \text{(by deduction)}\} \end{aligned}$$

Ontologies in Description Logics

Definition (Description logics (DLs))

Logics for use in knowledge representation with special attention on a good balance of expressibility and feasibility of reasoning services

- Can be mapped to fragments of FOL

Definition (Description logics (DLs))

Logics for use in knowledge representation with special attention on a good balance of expressibility and feasibility of reasoning services

- Can be mapped to fragments of FOL
- Usage
  - Ontology representation language
  - Foundation for standard web ontology language (OWL)

Definition (Description logics (DLs))

Logics for use in knowledge representation with special attention on a good balance of expressibility and feasibility of reasoning services

- Can be mapped to fragments of FOL
- Usage
  - Ontology representation language
  - Foundation for standard web ontology language (OWL)
- Have been investigated for ca. 30 years now
  - Many theoretical insights on various different purpose DLs
  - Various reasoners

Family of DLs

- Variable-free logics centered around concepts
- concepts = one-ary predicates in FOL = classes in OWL

Example (Concepts)

- $Students$ ("students")
- $Students \sqcap Male$ (" Male students")
- $\exists attends.MathCourse$ ("Those attending a math course")
- $\forall hasFriends.Freaks$ ("Those having only freaks as friends")
- $Person \sqcap \forall attends.(Course \sqcap \neg Easy)$ ("Persons attending only non-easy courses")

A Semi-Expressive Logic: $\mathcal{ALC}$

- Vocabulary: constants N_i, atomic concepts N_C, roles N_{Ro}

A Semi-Expressive Logic: $\mathcal{ALC}$

- Vocabulary: constants N_i, atomic concepts N_C, roles N_{Ro}
- Concepts: syntax

$$C \quad ::= \quad A \quad \text{for } A \in N_C \mid C \sqcap C \mid C \sqcup C \mid \neg C \mid \forall r.C \mid \exists r.C \text{ for } r \in N_{Ro} \mid \bot \mid \top$$

A Semi-Expressive Logic: $\mathcal{ALC}$

- Vocabulary: constants N_i, atomic concepts N_C, roles N_{Ro}
- Concepts: syntax

$$
\begin{aligned}
C \quad ::= \quad & A \quad \text{for } A \in N_C \mid C \sqcap C \mid C \sqcup C \mid \neg C \mid \\
& \forall r.C \mid \exists r.C \text{ for } r \in N_{Ro} \mid \bot \mid \top
\end{aligned}
$$

- Concepts: semantics

A Semi-Expressive Logic: $\mathcal{ALC}$

- Vocabulary: constants N_i, atomic concepts N_C, roles N_{Ro}
- Concepts: syntax

$$C \quad ::= \quad A \quad \text{for } A \in N_C \mid C \sqcap C \mid C \sqcup C \mid \neg C \mid \forall r.C \mid \exists r.C \text{ for } r \in N_{Ro} \mid \bot \mid \top$$

- Concepts: semantics

- Interpretation
  $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$
- $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$ for all $A \in N_C$
- $c^\mathcal{I} \in \Delta^\mathcal{I}$ for all $c \in N_i$
- $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$
  for all $r \in N_{Ro}$

A Semi-Expressive Logic: $\mathcal{ALC}$

- Vocabulary: constants N_i, atomic concepts N_C, roles N_{Ro}
- Concepts: syntax

$$C \quad ::= \quad A \quad \text{for } A \in N_C \mid C \sqcap C \mid C \sqcup C \mid \neg C \mid$$
$$\forall r.C \mid \exists r.C \text{ for } r \in N_{Ro} \mid \bot \mid \top$$

- Concepts: semantics

- Interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$
- $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ for all $A \in N_C$
- $c^{\mathcal{I}} \in \Delta^{\mathcal{I}}$ for all $c \in N_i$
- $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ for all $r \in N_{Ro}$

- $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$
- $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$
- $\neg C = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
- $(\forall r.C)^{\mathcal{I}} = \{d \in \Delta^{\mathcal{I}} \mid$ for all $e \in \Delta^{\mathcal{I}}$: If $(d, e) \in r^{\mathcal{I}}$ then $e \in C^{\mathcal{I}}\}$
- $(\exists r.C)^{\mathcal{I}} = \{d \in \Delta^{\mathcal{I}} \mid$ there is $e \in \Delta^{\mathcal{I}}$ s.t. $(d, e) \in r^{\mathcal{I}}$ and $e \in C^{\mathcal{I}}\}$

Tbox and Abox

- Terminological box (tbox) $\mathcal{T}$
  - Finite set of general concept inclusions (GCIs)

$$C \sqsubseteq D$$

  - Semantics: $\mathcal{I} \models C \sqsubseteq D$ iff $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

Tbox and Abox

- Terminological box (tbox) $\mathcal{T}$
  - Finite set of general concept inclusions (GCIs)

$$C \sqsubseteq D$$

  - Semantics: $\mathcal{I} \models C \sqsubseteq D$ iff $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

- Assertional box (abox) $\mathcal{A}$
  - Finite set of assertions
  - Assertion: $C(a)$, $r(a, b)$
  - Semantics:
    $\mathcal{I} \models C(a)$ iff $a^{\mathcal{I}} \in C^{\mathcal{I}}$
    $\mathcal{I} \models r(a, b)$ iff $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$.

Tbox and Abox

- Terminological box (tbox) $\mathcal{T}$
  - Finite set of general concept inclusions (GCIs)

$$C \sqsubseteq D$$

  - Semantics: $\mathcal{I} \models C \sqsubseteq D$ iff $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

- Assertional box (abox) $\mathcal{A}$
  - Finite set of assertions
  - Assertion: $C(a)$, $r(a, b)$
  - Semantics:
    $\mathcal{I} \models C(a)$ iff $a^{\mathcal{I}} \in C^{\mathcal{I}}$
    $\mathcal{I} \models r(a, b)$ iff $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$.

- Ontology: $(\sigma, \mathcal{T}, \mathcal{A})$

Definition (Basic Reasoning services)

- Model: $\mathcal{I} \models \mathcal{O}$ iff $\mathcal{I} \models \mathcal{T} \cup \mathcal{A}$ iff $\mathcal{I} \models ax$ for all $ax \in \mathcal{T} \cup \mathcal{A}$.
- Satisfiability/Consistency: $\mathcal{O}$ is satisfiable iff $\mathcal{T} \cup \mathcal{A}$ is satisfiable iff it has a model
- Entailment: $\mathcal{O} \vDash ax$ iff: For all $\mathcal{I}$: If $\mathcal{I} \models \mathcal{O}$ then $\mathcal{I} \models ax$.
- ...

Definition (Basic Reasoning services)

- Model: $\mathcal{I} \models \mathcal{O}$ iff $\mathcal{I} \models \mathcal{T} \cup \mathcal{A}$ iff $\mathcal{I} \models ax$ for all $ax \in \mathcal{T} \cup \mathcal{A}$.
- Satisfiability/Consistency: $\mathcal{O}$ is satisfiable iff $\mathcal{T} \cup \mathcal{A}$ is satisfiable iff it has a model
- Entailment: $\mathcal{O} \vDash ax$ iff: For all $\mathcal{I}$: If $\mathcal{I} \models \mathcal{O}$ then $\mathcal{I} \models ax$.
- ...

Definition (Extended Reasoning services)

- Query answering: Certain answers $cert(Q(x), \mathcal{O}) = \{\vec{a} \in N_i \mid \mathcal{O} \vDash Q[\vec{x}/\vec{a}]\}$
- ...

Example (Certain Answers for Conjunctive Queries)

$$\mathcal{T} = \{ \top \sqsubseteq \mathit{Male} \sqcup \mathit{Female}, \mathit{Male} \sqcap \mathit{Female} \sqsubseteq \bot \}$$

$$\mathcal{A} = \{ \mathit{friend}(\mathit{john}, \mathit{susan}), \mathit{friend}(\mathit{john}, \mathit{andrea}), \mathit{female}(\mathit{susan}), \mathit{likes}(\mathit{susan}, \mathit{andrea}), \mathit{likes}(\mathit{andrea}, \mathit{bill}), \mathit{Male}(\mathit{bill}) \}$$

$$Q(x) = \exists y, z(\mathit{friend}(x, y) \wedge \mathit{Female}(y) \wedge \mathit{likes}(y, z) \wedge \mathit{Male}(z))$$

- $\mathit{cert}(Q(x), \mathcal{O}) = ?$

Example (Certain Answers for Conjunctive Queries)

$$\begin{aligned}\mathcal{T} &= \{\ \top \sqsubseteq \mathit{Male} \sqcup \mathit{Female}, \mathit{Male} \sqcap \mathit{Female} \sqsubseteq \bot\ \} \\ \mathcal{A} &= \{\ \mathit{friend}(\mathit{john}, \mathit{susan}), \mathit{friend}(\mathit{john}, \mathit{andrea}), \mathit{female}(\mathit{susan}), \\ &\qquad \mathit{likes}(\mathit{susan}, \mathit{andrea}), \mathit{likes}(\mathit{andrea}, \mathit{bill}), \mathit{Male}(\mathit{bill})\ \}\end{aligned}$$

$$Q(x) = \exists y, z(\mathit{friend}(x, y) \wedge \mathit{Female}(y) \wedge \mathit{likes}(y, z) \wedge \mathit{Male}(z))$$

- $\mathit{cert}(Q(x), \mathcal{O}) =?$

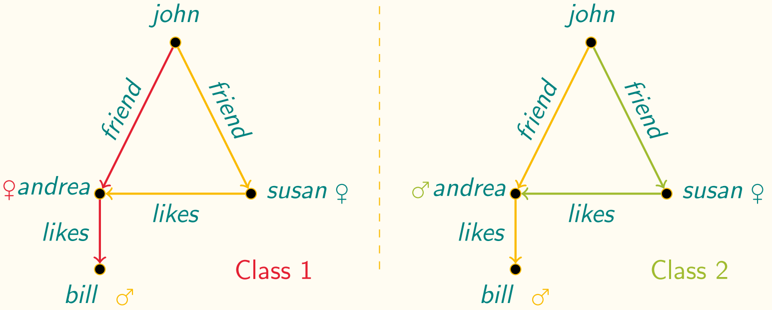
- We have to consider all possible models of the ontology
- But here there are actually two classes:
  Andrea is male vs. Andrea is not male.

Example (Certain Answers for Conjunctive Queries)

$$\mathcal{T} = \{ \top \sqsubseteq \mathit{Male} \sqcup \mathit{Female}, \mathit{Male} \sqcap \mathit{Female} \sqsubseteq \bot \}$$

$$\mathcal{A} = \{ \mathit{friend}(\mathit{john}, \mathit{susan}), \mathit{friend}(\mathit{john}, \mathit{andrea}), \mathit{female}(\mathit{susan}), \mathit{likes}(\mathit{susan}, \mathit{andrea}), \mathit{likes}(\mathit{andrea}, \mathit{bill}), \mathit{Male}(\mathit{bill}) \}$$

$$Q(x) = \exists y, z(\mathit{friend}(x, y) \land \mathit{Female}(y) \land \mathit{likes}(y, z) \land \mathit{Male}(z))$$



$$\mathit{cert}(Q(x), \mathcal{O}) = \{\mathit{john}\}$$

Cone-Embedding for $\mathcal{ALC}$ Ontologies

Aim

Find ML-feasible geometric models such that $\mathcal{ALC}$ ontology is classically satisfiable iff it is satisfiable by such a geometric model.

Simple Case: Propositional $\mathcal{ALC}$

- Boolean concepts

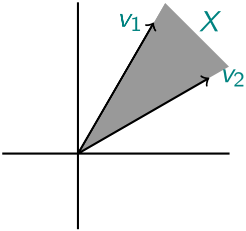
$$C \longrightarrow A \mid \bot \mid \top \mid \neg C \mid C \sqcap C \mid C \sqcup C$$

- Tbox restricted to Boolean concepts
- Geometric model for Boolean concepts based on closed convex cones
- Embedding space
  - $E = \mathbb{R}^n$ for some $n \in \mathbb{N}$
  - Scalar product for $v, w \in E$

$$\langle v, w \rangle = \sum_{1 \leq i \leq n} v_i w_i$$

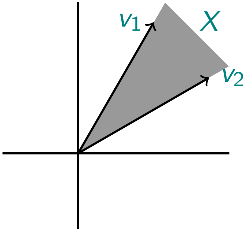
Cones

- $X \subseteq E$ is a convex cone iff for all $v, w \in X, \lambda, \mu \in \mathbb{R}_{\geq 0}$: $\lambda v + \mu w \in X$



Cones

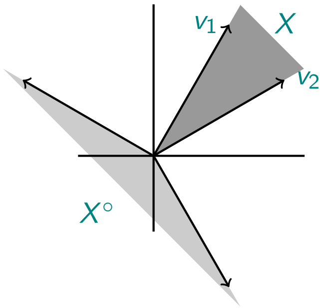
- $X \subseteq E$ is a convex cone iff for all $v, w \in X, \lambda, \mu \in \mathbb{R}_{\geq 0}$: $\lambda v + \mu w \in X$
- $conicHull(X)$ = smallest cone containing X



Cones

- $X \subseteq E$ is a convex cone iff for all $v, w \in X, \lambda, \mu \in \mathbb{R}_{\geq 0}$: $\lambda v + \mu w \in X$
- $conicHull(X)$ = smallest cone containing X
- Polar cone

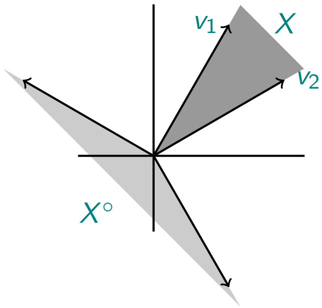
$$X^\circ = \{v \in E \mid \forall w \in X : \langle v, w \rangle \leq 0\}$$



Cones

- $X \subseteq E$ is a convex cone iff for all $v, w \in X, \lambda, \mu \in \mathbb{R}_{\geq 0}$: $\lambda v + \mu w \in X$
- $conicHull(X)$ = smallest cone containing X
- Polar cone

$$X^\circ = \{v \in E \mid \forall w \in X : \langle v, w \rangle \leq 0\}$$



Proposition

For closed convex cones X, Y:

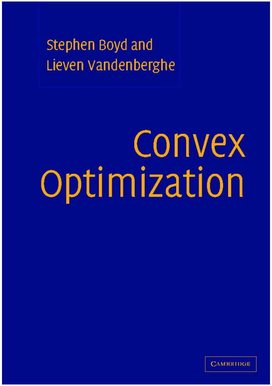
- X° *is a closed convex cone*
- $(X^\circ)^\circ = X$
- $conicHull(X \cup Y) = (X^\circ \cap Y^\circ)^\circ$

Definition (Cone-Semantics of $\mathcal{ALC}$ concepts (First Try))

- A geometric interpretation $\mathcal{I}$ interprets all atomic concepts by closed convex cones in E
- Semantics for arbitrary concepts under $\mathcal{I}$
  - $(\top)^{\mathcal{I}} = \Delta$
  - $(\bot)^{\mathcal{I}} = \{\vec{0}\}$
  - $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$
  - $(\neg C)^{\mathcal{I}} = C^{\circ}$
  - $(C \sqcup D)^{\mathcal{I}} = (\neg(\neg C \sqcap \neg D))^{\mathcal{I}}$

Why Convex Cones?

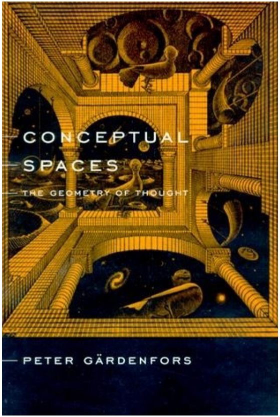
1. ML-feasibility:
   convex (even conic) optimization



Why Convex Cones?

1. ML-feasibility:
   convex (even conic) optimization

2. Linguistical/cognitive justification:
   conceptual spaces



Why Convex Cones?

1. ML-feasibility:
   convex (even conic) optimization

2. Linguistical/cognitive justification:
   conceptual spaces

3. Logic: polarity as negation

Lemma (Farkas' Lemma)

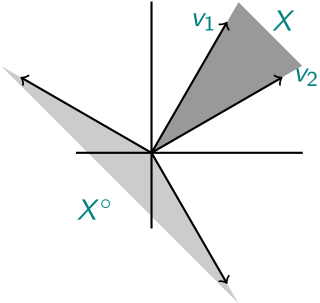
Let X be the convex cone generated by vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ and consider $w \in \mathbb{R}^n$. Then either $w \in X$ or there is $z \in \mathbb{R}^n$ such that $\langle z, v_i \rangle \leq 0$ for all i, and $\langle z, w \rangle \geq 0$.

Why Convex Cones?

1. ML-feasibility:
   convex (even conic) optimization

2. Linguistical/cognitive justification:
   conceptual spaces

3. Logic: polarity as negation



Lemma (Farkas' Lemma)

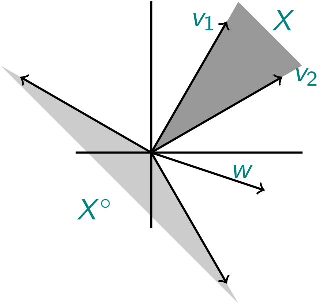
Let X be the convex cone generated by vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ and consider $w \in \mathbb{R}^n$. Then either $w \in X$ or there is $z \in \mathbb{R}^n$ such that $\langle z, v_i \rangle \leq 0$ for all i, and $\langle z, w \rangle \geq 0$.

Why Convex Cones?

1. ML-feasibility:
   convex (even conic) optimization

2. Linguistical/cognitive justification:
   conceptual spaces

3. Logic: polarity as negation



Lemma (Farkas' Lemma)

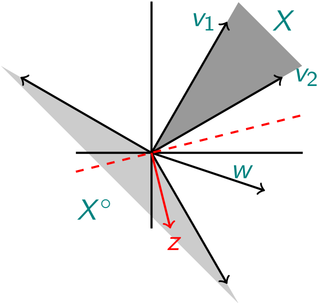
Let X be the convex cone generated by vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ and consider $w \in \mathbb{R}^n$. Then either $w \in X$ or there is $z \in \mathbb{R}^n$ such that $\langle z, v_i \rangle \leq 0$ for all i, and $\langle z, w \rangle \geq 0$.

Why Convex Cones?

1. ML-feasibility:
   convex (even conic) optimization

2. Linguistical/cognitive justification:
   conceptual spaces

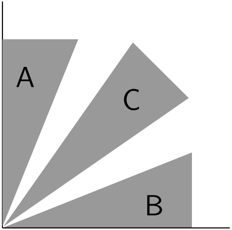
3. Logic: polarity as negation



Lemma (Farkas' Lemma)

Let X be the convex cone generated by vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ and consider $w \in \mathbb{R}^n$. Then either $w \in X$ or there is $z \in \mathbb{R}^n$ such that $\langle z, v_i \rangle \leq 0$ for all i, and $\langle z, w \rangle \geq 0$.

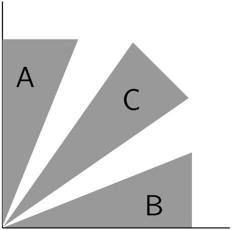
Cones don't like $\mathcal{ALC}$



$$(C \sqcap (A \sqcup B))^{\mathcal{I}} = C^{\mathcal{I}} \neq$$
$$((C \sqcap A) \sqcup (C \sqcap B))^{\mathcal{I}} = \bot^{\mathcal{I}}$$

- Distributivity law not fulfilled
- What should we do?

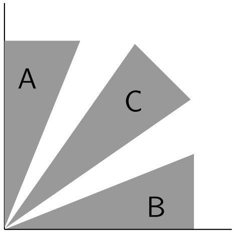
Cones don't like $\mathcal{ALC}$



$$(C \sqcap (A \sqcup B))^{\mathcal{I}} = C^{\mathcal{I}} \neq$$
$$((C \sqcap A) \sqcup (C \sqcap B))^{\mathcal{I}} = \bot^{\mathcal{I}}$$

- Distributivity law not fulfilled
- What should we do?
  1. Restrict cones (today)

Cones don't like $\mathcal{ALC}$



$$(C \sqcap (A \sqcup B))^{\mathcal{I}} = C^{\mathcal{I}} \neq$$
$$((C \sqcap A) \sqcup (C \sqcap B))^{\mathcal{I}} = \bot^{\mathcal{I}}$$

- Distributivity law not fulfilled
- What should we do?
  1. Restrict cones (today)
  2. Search for a (the) logic of cones (not today)

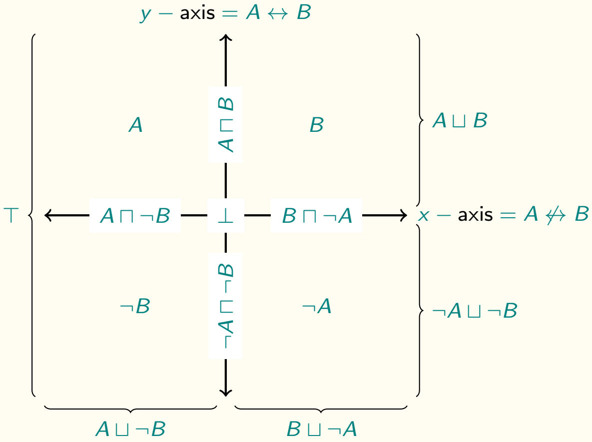
Searching for $\mathcal{ALC}$-cones (Nomen est Omen)

Definition

X is an al-cone in $\mathbb{R}^n$ iff $X = X_1 \times \cdots \times X_n$

where each $X_i \in \{\mathbb{R}, \mathbb{R}_+, \mathbb{R}_-, \{0\}\}$

Example (Boolean Algebra over Two Atomic Concepts)



Al-Cones Do the Job

Proposition

Boolean $\mathcal{ALC}$-ontologies are classically satisfiable iff they are by a geometric model over $\mathbb{R}^n$ based on al-cones of the form $X_1 \times \cdots \times X_n$ with $X_i \in \{\{0\}, \mathbb{R}^+, \mathbb{R}^-, \mathbb{R}\}$ for $i \in \{1, \ldots, n\}$.

- Consequence: Any learning method based on al-cones as hypotheses will succeed if learned positions consistent with ontology

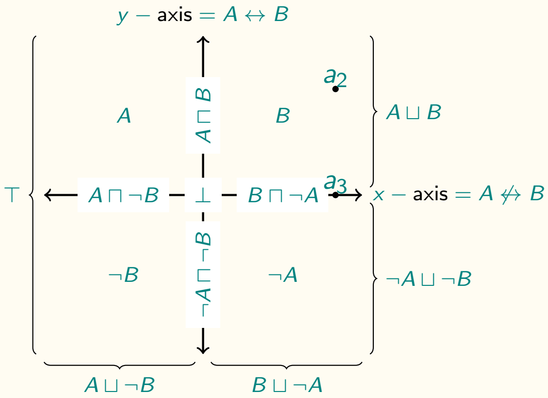
Al-Cones Do the Job

Proposition

Boolean $\mathcal{ALC}$-ontologies are classically satisfiable iff they are by a geometric model over $\mathbb{R}^n$ based on al-cones of the form $X_1 \times \cdots \times X_n$ with $X_i \in \{\{0\}, \mathbb{R}^+, \mathbb{R}^-, \mathbb{R}\}$ for $i \in \{1, \ldots, n\}$.

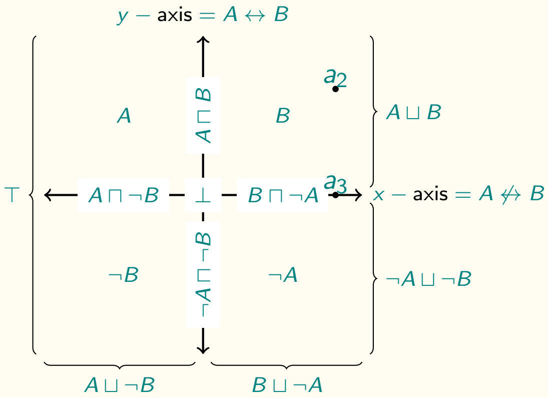
- Consequence: Any learning method based on al-cones as hypotheses will succeed if learned positions consistent with ontology
- Hypothesis space $\mathcal{H} = \{$ rotation(α) of al-cone system $\}$

Example



- Our geometric models are partial models

Example



- Our geometric models are partial models
- Construction before not faithful: any $c \in A \sqcup B$ must be A or B, though ontology does not say so.

Faithful Models

Proposition

For classically satisfiable Boolean $\mathcal{ALC}$-ontologies there is a concept-faithful and tbox-faithful geometric model on some $\mathbb{R}^{2n}$ based on al-cones of the form $X_1 \times \cdots \times X_{2n}$ with $X_{2i} \in \{\{0\}, \mathbb{R}^+, \mathbb{R}^-, \mathbb{R}\}$ and $X_{2i+1} = X_{2i}$.

Al-Cone Models for Full $\mathcal{ALC}$

- Previous construction does not work because tbox-induced Boolean algebra not atomic
- Approximative solution: consider rank bound
- Roles are interpreted "classically"

Proposition

$\mathcal{ALC}$-ontologies are classically satisfiable iff they are satisfiable by a (abox faithful, m-rank-concepts faithful) geometric model on some finite $\mathbb{R}^n$ using sets of the form $X_1 \times \cdots \times X_n$ with $X_i \in \{\{0\}, \mathbb{R}^+, \mathbb{R}^-, \mathbb{R}\}$.

Embeddings for Ontologies in Other Logics

Embedding for Datalog$^{\pm}$

- Existential rules (with atoms B_i, H_j)

$$B_1 \wedge \cdots \wedge B_n \rightarrow \exists X_1, \ldots, X_j . H_1 \wedge \cdots \wedge H_k$$

Example

$$\mathit{short\text{-}talk}(X) \wedge \mathit{well\text{-}planed}(X) \rightarrow \exists t . \mathit{lasts}(X, t) \wedge t < 1h$$

Embedding for Datalog$^{\pm}$

- Existential rules (with atoms B_i, H_j)

$$B_1 \wedge \cdots \wedge B_n \rightarrow \exists X_1, \ldots, X_j . H_1 \wedge \cdots \wedge H_k$$

Example

$$\textit{short-talk}(X) \wedge \textit{well-planed}(X) \rightarrow \exists t . \textit{lasts}(X, t) \wedge t < 1h$$

- Integrity constraints

$$B_1 \wedge \cdots \wedge B_n \rightarrow \bot$$

Example

$$\textit{short-talk}(X) \wedge \textit{long-talk}(X) \rightarrow \bot$$

Embedding for Datalog$^\pm$

- Existential rules (with atoms B_i, H_j)

$$B_1 \wedge \dots \wedge B_n \rightarrow \exists X_1, \dots, X_j . H_1 \wedge \dots \wedge H_k$$

Example

$$\mathit{short\text{-}talk}(X) \wedge \mathit{well\text{-}planed}(X) \rightarrow \exists t.\mathit{lasts}(X, t) \wedge t < 1h$$

- Integrity constraints

$$B_1 \wedge \dots \wedge B_n \rightarrow \bot$$

Example

$$\mathit{short\text{-}talk}(X) \wedge \mathit{long\text{-}talk}(X) \rightarrow \bot$$

- Integrity constraints capture only one aspect of negation (disjointness), but not covering (no negation on lhs)

Geometric Models Based on Convex Sets

Theorem (Basulto/Schockaert 18)

For all Datalog$^{\pm}$ KBs $\mathcal{O}$ with quasi-chained rules only:
$\mathcal{O}$ is classically satisfiable iff it is satisfiable by a geometric model that interprets all relations by convex sets.

$B_1 \wedge \dots \wedge B_n \rightarrow \exists X_1, \dots, X_j . H_1 \wedge \dots \wedge H_k$ is quasi-chained iff:

$$\mid (var(B_1) \cup \dots \cup var(B_{i-1})) \cap var(B_i) \mid \leq 1$$

Summary

- KG embedding as means to resolve trade-off

  ML-feasible geometric structures vs. logical semantics

Summary

- KG embedding as means to resolve trade-off
  ML-feasible geometric structures vs. logical semantics

- Step towards explainable AI

Summary

- KG embedding as means to resolve trade-off
  ML-feasible geometric structures vs. logical semantics

- Step towards explainable AI

Thanks for your attention!

References

Nickel, M.; Tresp, V.; and Kriegel, H.-P. A three- way model for collective learning on multi-relational data. In Proc. ICML, 809–816, 2011.

A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. Translating embeddings for modeling multi-relational data. In Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems, pp. 2787–2795, 2013.

B. Yang, W. tau Yih, X. He, J. Gao, and L. Deng. Embedding entities and relations for learning and inference in knowledge bases. CoRR, abs/1412.6575, 2014.

Nguyen, D. Q.; Sirts, K.; Qu, L.; and Johnson, M. STransE: a novel embedding model of entities and relationships in knowledge bases. In Proc. of NAACL-HLT, 460–466, 2016.

Trouillon, T., Welbl, J., Riedel, S., Gaussier, E. and Bouchard, G. Complex embeddings for simple link prediction. In Proc. ICML, 2071–2080, 2016.

S. Mehran Kazemi and D. Poole. SimplE Embedding for Link Prediction in Knowledge Graphs. arXiv e-prints, page arXiv:1802.04868, Feb 2018.

V. Gutierrez-Basulto and S. Schockaert. From knowledge graph embedding to ontology embedding? An analysis of the compatibility between vector space representations and rules. In Proceedings of KR 2018, 379–388, 2018.

Acknowledgements

- The presentation of KG embeddings with scoring functions follows that of (Gutierrez-Basulto/Schockaert 18) (see References)
- ESSLLI 2010 Course by Calvanese and Zakharyaschev
  http://www.inf.unibz.it/~calvanese/teaching/2010-08-ESSLLI-DL-QA/
- Reasoning Web Summer School 2014 course by Kontchakov on Description Logics
  http://rw2014.di.uoa.gr/sites/default/files/slides/An_Introduction_to_Description_Logics.pdf
- Lecture notes by Calvanese in 2013/2014 course on Ontology and Database Systems
  https://www.inf.unibz.it/~calvanese/teaching/14-15-odbs/lecture-notes/
- Course notes by Franz Baader on Description Logics
  https://lat.inf.tu-dresden.de/teaching/ws2017-2018/DL/
- Pictures on RESCAL copyright of Matthias Nickels