Mining Linguistic Associations from Data Using LFLC

Antonín Dvořák, Vilém Novák, Irina Perfilieva

University of Ostrava
Institute for Research and Applications of Fuzzy Modeling
30. dubna 22, 701 03 Ostrava 1, Czech Republic
antonin.dvorak@osu.cz

Data – Algorithms – Decision making, Třešť,
11. 12. 2006

Outline

Introduction

Method for finding linguistically characterized associations in large databases.

Example:

high *profit* and **rather low** *cost*

$\sim$

very high *productivity* and **significantly large** *volume of sale*

Characteristic feature — **evaluating linguistic expressions and predications**

Introduction

Method for finding linguistically characterized associations in large databases.

Example:

high *profit* and **rather low** *cost*

$\sim$

very high *productivity* and **significantly large** *volume of sale*

Characteristic feature — **evaluating linguistic expressions and predications**

Introduction

Method for finding linguistically characterized associations in large databases.

Example:

high *profit* and **rather low** *cost*

$\sim$

very high *productivity* and **significantly large** *volume of sale*

Characteristic feature — **evaluating linguistic expressions and predications**

Methods

Two methods:

- numbers replaced by **evaluating linguistic expressions**
  mining linguistic associations — **GUHA method**
  (P. Hájek, T. Havránek, 1968, 1978)
- technique of **fuzzy transform**

Methods

Two methods:

- numbers replaced by **evaluating linguistic expressions**
  mining linguistic associations — **GUHA method**
  (P. Hájek, T. Havránek, 1968, 1978)
- technique of **fuzzy transform**

Advantages:

- Easy (at least easier) understandability
- Use of logical properties for reduction of the number of associations
- Vague meaning enables less strict interpretation

Advantages:

- Easy (at least easier) understandability
- Use of logical properties for reduction of the number of associations
- Vague meaning enables less strict interpretation

Advantages:

- Easy (at least easier) understandability
- Use of logical properties for reduction of the number of associations
- Vague meaning enables less strict interpretation

Evaluating expressions

Logical theory of their meaning

- Atomic: *small, medium, big* (canonical words)
- Fuzzy quantities: *about twenty, roughly 100*
- Simple: *very small, more or less medium, roughly big, about thirty five, roughly one thousand*
- Compound: *very roughly small or medium*
- Fuzzy IF-THEN rules: conditional linguistic statements

Evaluating expressions

Logical theory of their meaning

- Atomic: *small, medium, big* (canonical words)
- Fuzzy quantities: *about twenty, roughly 100*
- Simple: *very small, more or less medium, roughly big, about thirty five, roughly one thousand*
- Compound: *very roughly small or medium*
- Fuzzy IF-THEN rules: conditional linguistic statements

Evaluating expressions

Logical theory of their meaning

- Atomic: *small, medium, big* (canonical words)
- Fuzzy quantities: *about twenty, roughly 100*
- Simple: *very small, more or less medium, roughly big, about thirty five, roughly one thousand*
- Compound: *very roughly small or medium*
- Fuzzy IF-THEN rules: conditional linguistic statements

Evaluating expressions

Logical theory of their meaning

- Atomic: *small, medium, big* (canonical words)
- Fuzzy quantities: *about twenty, roughly 100*
- Simple: *very small, more or less medium, roughly big, about thirty five, roughly one thousand*
- Compound: *very roughly small or medium*
- Fuzzy IF-THEN rules: conditional linguistic statements

Evaluating expressions

Logical theory of their meaning

- Atomic: *small, medium, big* (canonical words)
- Fuzzy quantities: *about twenty, roughly 100*
- Simple: *very small, more or less medium, roughly big, about thirty five, roughly one thousand*
- Compound: *very roughly small or medium*
- Fuzzy IF-THEN rules: conditional linguistic statements

Evaluating predications

Evaluating linguistic predication

⟨noun phrase⟩ is $\mathcal{A}$ or $\mathcal{A}$ ⟨noun phrase⟩

Example: ⟨ temperature of melted metal ⟩ is *very high*
very high ⟨ temperature of melted metal⟩

$$\mathcal{C} := \bigwedge_{i\in I}(\mathcal{A}_i\, X_i) \qquad \mathcal{D} := \bigvee_{i\in I}(\mathcal{B}_i\, X_i)$$

$$\mathcal{E} := \bigvee_{j\in J}\mathcal{C}_j \qquad \mathcal{F} := \bigwedge_{j\in J}\mathcal{D}_j$$

$\bigwedge$ — linguistic conjunction (“and”),
$\bigvee$ — linguistic disjunction (“or”)

Evaluating predications

Evaluating linguistic predication

$\langle$noun phrase$\rangle$ is $\mathcal{A}$ or $\mathcal{A}$ $\langle$noun phrase$\rangle$

Example: $\langle$ temperature of melted metal $\rangle$ is *very high*
very high $\langle$ temperature of melted metal$\rangle$

$$\mathcal{C} := \bigwedge_{i \in I} (\mathcal{A}_i\, X_i) \qquad \mathcal{D} := \bigvee_{i \in I} (\mathcal{B}_i\, X_i)$$

$$\mathcal{E} := \bigvee_{j \in J} \mathcal{C}_j \qquad \mathcal{F} := \bigwedge_{j \in J} \mathcal{D}_j$$

$\bigwedge$ — linguistic conjunction (“and”),
$\bigvee$ — linguistic disjunction (“or”)

Evaluating linguistic predication

⟨noun phrase⟩ is $\mathcal{A}$ or $\mathcal{A}$ ⟨noun phrase⟩

Example: ⟨ temperature of melted metal ⟩ is *very high*
very high ⟨ temperature of melted metal⟩

$$\mathcal{C} := \bigwedge_{i \in I} (\mathcal{A}_i \; X_i) \qquad \mathcal{D} := \bigvee_{i \in I} (\mathcal{B}_i \; X_i)$$

$$\mathcal{E} := \bigvee_{j \in J} \mathcal{C}_j \qquad \mathcal{F} := \bigwedge_{j \in J} \mathcal{D}_j$$

$\bigwedge$ — linguistic conjunction (“and”),
$\bigvee$ — linguistic disjunction (“or”)

Mathematical model of the meaning of evaluating expressions

Context, intension, extension

- Intension of $\mathcal{A}$
$$A : W \longrightarrow \mathcal{F}(V).$$
- *Context:* $\langle v_L, v_S, v_R \rangle \mapsto [v_L, v_R]$
- *Extension* of $\mathcal{A}$ in a context $w \in W$ is a *fuzzy set* $A(w) \underset{\sim}{\subset} V$

Mathematical model of the meaning of evaluating expressions

Context, intension, extension

- Intension of $\mathcal{A}$

$$A : W \longrightarrow \mathcal{F}(V).$$

- *Context:* $\langle v_L, v_S, v_R \rangle \mapsto [v_L, v_R]$
- *Extension* of $\mathcal{A}$ in a context $w \in W$ is a *fuzzy set* $A(w) \underset{\sim}{\subset} V$

Associations Mining
Dvořák, Novák, Perfilieva
Introduction
Evaluating expressions
Linguistic associations
GUHA quantifiers
Rule reduction
FT
Associations using FT
LFLC
Experiments
Linguistic associations
NO_2 FT associations
Future work

Mathematical model of the meaning of evaluating expressions

Context, intension, extension

- Intension of $\mathcal{A}$

$$A : W \longrightarrow \mathcal{F}(V).$$

- *Context:* $\langle v_L, v_S, v_R \rangle \mapsto [v_L, v_R]$
- *Extension* of $\mathcal{A}$ in a context $w \in W$ is a *fuzzy set* $A(w) \underset{\sim}{\subset} V$

Mathematical model

Mathematical model of the meaning of evaluating expressions

Context, intension, extension

- Intension of $\mathcal{A}$

$$A : W \longrightarrow \mathcal{F}(V).$$

- *Context:* $\langle v_L, v_S, v_R \rangle \mapsto [v_L, v_R]$
- *Extension* of $\mathcal{A}$ in a context $w \in W$ is a *fuzzy set* $A(w) \underset{\sim}{\subset} V$

Hedges with *narrowing* and *widening* effect

Concrete hedges:

extremely (Ex), significantly (Si), very (Ve), empty hedge, more or less (ML), roughly (Ro), quite roughly (QR), very roughly (VR).

$$\text{Ex} \preceq \text{Si} \preceq \text{Ve} \preceq \langle\text{empty hedge}\rangle \preceq \text{ML} \preceq \text{Ro} \preceq \text{QR} \preceq \text{VR}$$

Induced specificity ordering of evaluating expressions

$$\langle\text{hedge}\rangle_1\langle\text{atomic term}\rangle \preceq \langle\text{hedge}\rangle_2\langle\text{atomic term}\rangle \quad \text{iff} \quad \langle\text{hedge}\rangle_1 \preceq \langle\text{hedge}\rangle_2$$

Linguistic hedges

Hedges with *narrowing* and *widening* effect

Concrete hedges:

extremely (Ex), significantly (Si), very (Ve), empty hedge, more or less (ML), roughly (Ro), quite roughly (QR), very roughly (VR).

$$\text{Ex} \preceq \text{Si} \preceq \text{Ve} \preceq \langle\text{empty hedge}\rangle \preceq \text{ML} \preceq \text{Ro} \preceq \text{QR} \preceq \text{VR}$$

Induced specificity ordering of evaluating expressions

$$\langle\text{hedge}\rangle_1\langle\text{atomic term}\rangle \preceq \langle\text{hedge}\rangle_2\langle\text{atomic term}\rangle \quad \text{iff} \quad \langle\text{hedge}\rangle_1 \preceq \langle\text{hedge}\rangle_2$$

Linguistic hedges

Hedges with *narrowing* and *widening* effect

Concrete hedges:

extremely (Ex), significantly (Si), very (Ve), empty hedge, more or less (ML), roughly (Ro), quite roughly (QR), very roughly (VR).

$$\text{Ex} \preceq \text{Si} \preceq \text{Ve} \preceq \langle\text{empty hedge}\rangle \preceq \text{ML} \preceq \text{Ro} \preceq \text{QR} \preceq \text{VR}$$

Induced specificity ordering of evaluating expressions

$$\langle\text{hedge}\rangle_1\langle\text{atomic term}\rangle \preceq \langle\text{hedge}\rangle_2\langle\text{atomic term}\rangle \quad \text{iff} \quad \langle\text{hedge}\rangle_1 \preceq \langle\text{hedge}\rangle_2$$

Linguistic hedges

Hedges with *narrowing* and *widening* effect

Concrete hedges:

extremely (Ex), significantly (Si), very (Ve), empty hedge, more or less (ML), roughly (Ro), quite roughly (QR), very roughly (VR).

$$\text{Ex} \preceq \text{Si} \preceq \text{Ve} \preceq \langle\text{empty hedge}\rangle \preceq \text{ML} \preceq \text{Ro} \preceq \text{QR} \preceq \text{VR}$$

Induced specificity ordering of evaluating expressions

$$\langle\text{hedge}\rangle_1\langle\text{atomic term}\rangle \preceq \langle\text{hedge}\rangle_2\langle\text{atomic term}\rangle \quad \text{iff} \quad \langle\text{hedge}\rangle_1 \preceq \langle\text{hedge}\rangle_2$$

Finding a suitable expression

Given an element $u \in w$, transform it into a suitable *perception*

$$\text{Suit} : \langle u, w \rangle \mapsto \mathcal{A}$$

Suit(u, w) gives (intension of) an evaluating expression $\mathcal{A}$ such that the observation $u \in w$ is the *most specific and typical* for extension of $\mathcal{A}$ in the context w

Finding a suitable expression

Given an element $u \in w$, transform it into a suitable *perception*

$$\text{Suit} : \langle u, w \rangle \mapsto \mathcal{A}$$

Suit(u, w) gives (intension of) an evaluating expression $\mathcal{A}$ such that the observation $u \in w$ is the *most specific and typical* for extension of $\mathcal{A}$ in the context w

Data table

Mining linguistic knowledge from data

Data

	X_1	$\cdots$	X_i	$\cdots$	X_n
o_1	$e_{X_1}(o_1)$	$\cdots$	$e_{X_i}(o_1)$	$\cdots$	$e_{X_n}(o_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_j	$e_{X_1}(o_j)$	$\cdots$	$e_{X_i}(o_j)$	$\cdots$	$e_{X_n}(o_j)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_m	$e_{X_1}(o_m)$	$\cdots$	$e_{X_i}(o_m)$	$\cdots$	$e_{X_n}(o_m)$

$f_{ji} \in \mathbb{R}$.

Mining pure linguistic associations

$$Ev_{ji} = \text{Suit}(e_{X_i}(o_j), w_i).$$

Convert the data into linguistic form

	X_1	$\cdots$	X_i	$\cdots$	X_n
o_1	Ev_{11}	$\cdots$	Ev_{1i}	$\cdots$	Ev_{1n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_j	Ev_{j1}	$\cdots$	Ev_{ji}	$\cdots$	Ev_{jn}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o'_m	$Ev_{m'1}$	$\cdots$	$Ev_{m'i}$	$\cdots$	$Ev_{m'n}$

Generally smaller size ($m' \ll m$)

Mining pure linguistic associations

$$Ev_{ji} = \text{Suit}(e_{X_i}(o_j), w_i).$$

Convert the data into linguistic form

	X_1	$\cdots$	X_i	$\cdots$	X_n
o_1	Ev_{11}	$\cdots$	Ev_{1i}	$\cdots$	Ev_{1n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_j	Ev_{j1}	$\cdots$	Ev_{ji}	$\cdots$	Ev_{jn}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o'_m	$Ev_{m'1}$	$\cdots$	$Ev_{m'i}$	$\cdots$	$Ev_{m'n}$

Generally smaller size ($m' \ll m$)

Linguistic associations

$$\underbrace{\bigwedge_{i=1}^{p}(\mathcal{A}_i\ Y_i)}_{\mathcal{C}} \sim \underbrace{\bigwedge_{j=1}^{q}(\mathcal{B}_j\ Z_j)}_{\mathcal{D}}$$

After being assigned, linguistic predications $\mathcal{C}$, $\mathcal{D}$ behave as logical data

For each object o_j, it is true (or not true) that the attribute X_i is evaluated by the expression $\mathcal{A}_{ji}$

Suit acts as special partition operator

Linguistic associations

$$\underbrace{\bigwedge_{i=1}^{p}(\mathcal{A}_i\ Y_i)}_{\mathcal{C}} \sim \underbrace{\bigwedge_{j=1}^{q}(\mathcal{B}_j\ Z_j)}_{\mathcal{D}}$$

After being assigned, linguistic predications $\mathcal{C}$, $\mathcal{D}$ behave as logical data

For each object o_j, it is true (or not true) that the attribute X_i is evaluated by the expression $\mathcal{A}_{ji}$

Suit acts as special partition operator

GUHA quantifiers

Apply the standard GUHA quantifiers for mining associations

four-fold table

	$\mathcal{D}$	not $\mathcal{D}$
$\mathcal{C}$	a	b
not $\mathcal{C}$	c	d

(1)

- $\sqsubset_r^\gamma$ — binary multitudinal quantifier
  true, if $a > \gamma(a + b)$ and $a > r$
  γ – degree of confidence, r – degree of support
- $\sim_x$ — symmetric associational quantifier
  true if $ad > bc$

GUHA quantifiers

Apply the standard GUHA quantifiers for mining associations

four-fold table

	$\mathcal{D}$	not $\mathcal{D}$
$\mathcal{C}$	a	b
not $\mathcal{C}$	c	d

(1)

- $\sqsubset_r^\gamma$ — binary multitudinal quantifier
  true, if $a > \gamma(a+b)$ and $a > r$
  γ – degree of confidence, r – degree of support
- $\sim_x$ — symmetric associational quantifier
  true if $ad > bc$

Apply the standard GUHA quantifiers for mining associations

four-fold table

$$\begin{array}{c|c|c|} & \mathcal{D} & \text{not } \mathcal{D} \\ \hline \mathcal{C} & a & b \\ \hline \text{not } \mathcal{C} & c & d \\ \hline \end{array} \tag{1}$$

- $\sqsubset_r^\gamma$ — binary multitudinal quantifier
  true, if $a > \gamma(a + b)$ and $a > r$
  γ – degree of confidence, r – degree of support
- $\sim_x$ — symmetric associational quantifier
  true if $ad > bc$

Linguistic associations (Cont.)

Linguistic associations

$$\mathcal{C} \sqsubset_r^\gamma \mathcal{D}$$

hypotheses about possible validity of fuzzy IF-THEN rules

$$\mathcal{R} := \text{ IF } \mathcal{C} \text{ THEN } \mathcal{D}$$

Linguistic associations (Cont.)

Linguistic associations

$$\mathcal{C} \sqsubset_r^{\gamma} \mathcal{D}$$

hypotheses about possible validity of fuzzy IF-THEN rules

$$\mathcal{R} := \text{ IF } \mathcal{C} \text{ THEN } \mathcal{D}$$

Syntactic entailment

Reduction of number of mined linguistic associations

K — a set of mined linguistic associations

mining from the shortest and narrowest conjunctions

Syntactic entailment

if $\mathcal{A} \sqsubset_r^\gamma \mathcal{B}$ implies $\mathcal{C} \sqsubset_r^\gamma \mathcal{D}$ then

$$(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D})$$

Reduction of number of mined linguistic associations

K — a set of mined linguistic associations

mining from the shortest and narrowest conjunctions

Syntactic entailment

if $\mathcal{A} \sqsubset_r^{\gamma} \mathcal{B}$ implies $\mathcal{C} \sqsubset_r^{\gamma} \mathcal{D}$ then

$$(\mathcal{A} \sqsubset_r^{\gamma} \mathcal{B}) \vdash (\mathcal{C} \sqsubset_r^{\gamma} \mathcal{D})$$

Syntactic entailment

Reduction of number of mined linguistic associations

K — a set of mined linguistic associations

mining from the shortest and narrowest conjunctions

Syntactic entailment

if $\mathcal{A} \sqsubset_r^\gamma \mathcal{B}$ implies $\mathcal{C} \sqsubset_r^\gamma \mathcal{D}$ then

$$(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D})$$

Syntactic Entailment

Theorem

Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ be a linguistic predications.

(a) *If $\mathcal{D} \preceq \mathcal{D}'$ then $(\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}')$*

Example: (*big X* $\sqsubset_r^\gamma$ *small Y*) $\vdash$ (*big X* $\sqsubset_r^\gamma$ *roughly small Y*)

(b) $(\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}\ \mathrm{OR}\ \mathcal{B})$

Example: (*big X* $\sqsubset_r^\gamma$ *small Y*) $\vdash$ (*big X* $\sqsubset_r^\gamma$ *small Y* OR *medium Y*)

(c) $(\mathcal{A} \sqsubset_r^\gamma \mathcal{C},\ \mathcal{B} \sqsubset_r^\gamma \mathcal{C},\ \mathcal{A}\ \mathrm{AND}\ \mathcal{B} \sqsubset_s^\gamma \mathcal{C}) \vdash (\mathcal{A}\ \mathrm{OR}\ \mathcal{B} \sqsubset_r^\gamma \mathcal{C})$, *where $s \leq r$*

Syntactic Entailment

Theorem

Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ be a linguistic predications.

(a) *If $\mathcal{D} \preceq \mathcal{D}'$ then* $(\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}')$

Example: (*big X* $\sqsubset_r^\gamma$ *small Y*) $\vdash$ (*big X* $\sqsubset_r^\gamma$ *roughly small Y*)

(b) $(\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}\ \mathrm{OR}\ \mathcal{B})$

Example: (*big X* $\sqsubset_r^\gamma$ *small Y*) $\vdash$ (*big X* $\sqsubset_r^\gamma$ *small Y* OR *medium Y*)

(c) $(\mathcal{A} \sqsubset_r^\gamma \mathcal{C},\ \mathcal{B} \sqsubset_r^\gamma \mathcal{C},\ \mathcal{A}\ \mathrm{AND}\ \mathcal{B} \sqsubset_s^\gamma \mathcal{C}) \vdash (\mathcal{A}\ \mathrm{OR}\ \mathcal{B} \sqsubset_r^\gamma \mathcal{C})$, *where* $s \leq r$

Syntactic Entailment

Theorem

Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ be a linguistic predications.

(a) *If $\mathcal{D} \preceq \mathcal{D}'$ then $(\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}')$*

Example: $(big\ X \sqsubset_r^\gamma small\ Y) \vdash (big\ X \sqsubset_r^\gamma roughly\ small\ Y)$

(b) $(\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \vdash (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}\ \mathrm{OR}\ \mathcal{B})$

Example: $(big\ X \sqsubset_r^\gamma small\ Y) \vdash (big\ X \sqsubset_r^\gamma small\ Y\ \mathrm{OR}\ medium\ Y)$

(c) $(\mathcal{A} \sqsubset_r^\gamma \mathcal{C},\ \mathcal{B} \sqsubset_r^\gamma \mathcal{C},\ \mathcal{A}\ \mathrm{AND}\ \mathcal{B} \sqsubset_s^\gamma \mathcal{C}) \vdash (\mathcal{A}\ \mathrm{OR}\ \mathcal{B} \sqsubset_r^\gamma \mathcal{C})$, *where $s \leq r$*

Semantic entailment

Let $H_1, H_2 \subset K$ — two sets of mined associations

$$H_1 \models H_2.$$

Associations from H_1 are *more informative* than those from H_2 (the latter are *less informative* than the former)

1. **Rule of strong entailment**
   If $(\mathcal{A} \sim \mathcal{B}) \vdash (\mathcal{C} \sim \mathcal{D})$ then $(\mathcal{A} \sim \mathcal{B}) \models (\mathcal{C} \sim \mathcal{D})$.
2. **Rule of specificity**
   Let $(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}), (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \in K, \quad \mathcal{C} \preceq \mathcal{A}$ and $\mathcal{B} \preceq \mathcal{D}$.
   Then $(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}) \models (\mathcal{C} \sqsubset_r^\gamma \mathcal{D})$

   Example:
   $(\textit{big } X \sqsubset_r^\gamma \textit{ small } Y), (\textit{very big } X \sqsubset_r^\gamma \textit{ roughly small } Y) \in K$
   Then $(\textit{big } X \sqsubset_r^\gamma \textit{ small } Y) \models (\textit{very big } X \sqsubset_r^\gamma \textit{ roughly small } Y)$

Semantic entailment

Let $H_1, H_2 \subset K$ — two sets of mined associations

$$H_1 \models H_2.$$

Associations from H_1 are *more informative* than those from H_2 (the latter are *less informative* than the former)

1. **Rule of strong entailment**
   If $(\mathcal{A} \sim \mathcal{B}) \vdash (\mathcal{C} \sim \mathcal{D})$ then $(\mathcal{A} \sim \mathcal{B}) \models (\mathcal{C} \sim \mathcal{D})$.

2. **Rule of specificity**
   Let $(\mathcal{A} \sqsubset_r^{\gamma} \mathcal{B}), (\mathcal{C} \sqsubset_r^{\gamma} \mathcal{D}) \in K, \quad \mathcal{C} \preceq \mathcal{A}$ and $\mathcal{B} \preceq \mathcal{D}$.
   Then $(\mathcal{A} \sqsubset_r^{\gamma} \mathcal{B}) \models (\mathcal{C} \sqsubset_r^{\gamma} \mathcal{D})$

   Example:
   $(\textit{big } X \sqsubset_r^{\gamma} \textit{small } Y), (\textit{very big } X \sqsubset_r^{\gamma} \textit{roughly small } Y) \in K$
   Then $(\textit{big } X \sqsubset_r^{\gamma} \textit{small } Y) \models (\textit{very big } X \sqsubset_r^{\gamma} \textit{roughly small } Y)$

Semantic entailment

Let $H_1, H_2 \subset K$ — two sets of mined associations

$$H_1 \models H_2.$$

Associations from H_1 are *more informative* than those from H_2 (the latter are *less informative* than the former)

1. **Rule of strong entailment**
   If $(\mathcal{A} \sim \mathcal{B}) \vdash (\mathcal{C} \sim \mathcal{D})$ then $(\mathcal{A} \sim \mathcal{B}) \models (\mathcal{C} \sim \mathcal{D})$.
2. **Rule of specificity**
   Let $(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}), (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \in K, \quad \mathcal{C} \preceq \mathcal{A}$ and $\mathcal{B} \preceq \mathcal{D}$.
   Then $(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}) \models (\mathcal{C} \sqsubset_r^\gamma \mathcal{D})$

   Example:
   $(big\ X \sqsubset_r^\gamma small\ Y), (very\ big\ X \sqsubset_r^\gamma roughly\ small\ Y) \in K$
   Then $(big\ X \sqsubset_r^\gamma small\ Y) \models (very\ big\ X \sqsubset_r^\gamma roughly\ small\ Y)$

Semantic entailment

Let $H_1, H_2 \subset K$ — two sets of mined associations

$$H_1 \models H_2.$$

Associations from H_1 are *more informative* than those from H_2 (the latter are *less informative* than the former)

1. **Rule of strong entailment**
   If $(\mathcal{A} \sim \mathcal{B}) \vdash (\mathcal{C} \sim \mathcal{D})$ then $(\mathcal{A} \sim \mathcal{B}) \models (\mathcal{C} \sim \mathcal{D})$.
2. **Rule of specificity**
   Let $(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}), (\mathcal{C} \sqsubset_r^\gamma \mathcal{D}) \in K, \quad \mathcal{C} \preceq \mathcal{A}$ and $\mathcal{B} \preceq \mathcal{D}$.
   Then $(\mathcal{A} \sqsubset_r^\gamma \mathcal{B}) \models (\mathcal{C} \sqsubset_r^\gamma \mathcal{D})$

   Example:
   $(\textit{big } X \sqsubset_r^\gamma \textit{ small } Y), (\textit{very big } X \sqsubset_r^\gamma \textit{ roughly small } Y) \in K$
   Then $(\textit{big } X \sqsubset_r^\gamma \textit{ small } Y) \models (\textit{very big } X \sqsubset_r^\gamma \textit{ roughly small } Y)$

Semantic entailment (Cont.)

3 **Rule of disjunction**

Let $H = \{\mathcal{A}_j \sqsubset_r^\gamma \mathcal{C} \mid j \in J\} \subset K$, $\mathcal{B} :=$ $\mathrm{OR}_{j\in J}\ \mathcal{A}_j$ and $\mathcal{B} \sqsubset_r^\gamma \mathcal{C} \in K$. Then

(a) $\mathcal{B} \sqsubset_r^\gamma \mathcal{C} \models H$,

(b) $H \models \mathcal{B} \sqsubset_r^\gamma \mathcal{C}$.

Example:

$\{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y), (small\ X \text{ OR } medium\ X \sqsubset_r^\gamma big\ Y))\} \subset K$

Then

$$(small\ X \text{ OR } medium\ X \sqsubset_r^\gamma big\ Y)) \models \{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y)\},$$

$$\{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y)\} \models (small\ X \text{ OR } medium\ X \sqsubset_r^\gamma big\ Y))$$

Semantic entailment (Cont.)

3. **Rule of disjunction**
Let $H = \{\mathcal{A}_j \sqsubset_r^\gamma \mathcal{C} \mid j \in J\} \subset K$, $\mathcal{B} := \text{OR}_{j\in J}\ \mathcal{A}_j$ and $\mathcal{B} \sqsubset_r^\gamma \mathcal{C} \in K$. Then
(a) $\mathcal{B} \sqsubset_r^\gamma \mathcal{C} \models H$,
(b) $H \models \mathcal{B} \sqsubset_r^\gamma \mathcal{C}$.

Example:
$\{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y), (small\ X\ \text{OR}\ \ medium\ X \sqsubset_r^\gamma big\ Y)\} \subset K$

Then

$$(small\ X\ \text{OR}\ \ medium\ X \sqsubset_r^\gamma big\ Y)) \models \{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y)\},$$

$$\{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y)\} \models (small\ X\ \text{OR}\ \ medium\ X \sqsubset_r^\gamma big\ Y))$$

Semantic entailment (Cont.)

3. **Rule of disjunction**
Let $H = \{\mathcal{A}_j \sqsubset_r^\gamma \mathcal{C} \mid j \in J\} \subset K$, $\mathcal{B} := \text{OR}_{j\in J}\ \mathcal{A}_j$ and $\mathcal{B} \sqsubset_r^\gamma \mathcal{C} \in K$. Then
(a) $\mathcal{B} \sqsubset_r^\gamma \mathcal{C} \models H$,
(b) $H \models \mathcal{B} \sqsubset_r^\gamma \mathcal{C}$.

Example:
$\{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y),$
$(small\ X\ \text{OR}\ medium\ X \sqsubset_r^\gamma big\ Y))\} \subset K$

Then

$$(small\ X\ \text{OR}\ medium\ X \sqsubset_r^\gamma big\ Y)) \models \{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y)\},$$

$$\{(small\ X \sqsubset_r^\gamma big\ Y), (medium\ X \sqsubset_r^\gamma big\ Y)\} \models (small\ X\ \text{OR}\ medium\ X \sqsubset_r^\gamma big\ Y))$$

Semantic entailment (Cont.)

4 Rule of empty predication

$$\text{weak heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$
$$\text{roughly medium heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$
$$\text{more or less strong heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$

then

$$\text{heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$

is more informative

Semantic entailment (Cont.)

4 Rule of empty predication

$$\text{weak heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$
$$\text{roughly medium heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$
$$\text{more or less strong heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$

then

$$\text{heating} \sqsubset_r^\gamma \text{medium temperature of melted metal}$$

is more informative

Reduction

Reduction of the set K

If $H_1, H_2 \in K$ and $H_1 \models H_2$ then derive a new set

$$K' = K - H_2$$

Fuzzy transform

- Continuous function $f(x) : w \longrightarrow \mathbb{R}, \qquad w = [v_L, v_R]$
- $f(x)$ known at points $x_1, \dots, x_N$
- equidistant nodes $x_{0,1}, \dots, x_{0,n}$
- *n fuzzy numbers* Fn_{ν,x_0} (basic functions) — covering of w (extensions of “approximately x_0”)

Direct F-transform Values of $f(x)$ transformed into n-tuple of components $[F_1, \dots, F_n]$

$$F_k = \frac{\sum_{j=1}^{N} f(x_j)\,\mathrm{Fn}_{\nu,x_{0k}}(x_j)}{\sum_{j=1}^{N} \mathrm{Fn}_{\nu,x_{0k}}(x_j)}, \qquad k = 1, \dots, n.$$

Inverse F-transform Transform $[F_1, \dots, F_n]$ back

$$f_{F,n}(x) = \sum_{k=1}^{n} F_k \cdot \mathrm{Fn}_{\nu,x_{0k}}(x_j).$$

if n increases then $f_{F,n}(x_j)$ converges to $f(x_j)$

Data (again):

	X_1	$\cdots$	X_i	$\cdots$	X_n
o_1	$e_{X_1}(o_1)$	$\cdots$	$e_{X_i}(o_1)$	$\cdots$	$e_{X_n}(o_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_j	$e_{X_1}(o_j)$	$\cdots$	$e_{X_i}(o_j)$	$\cdots$	$e_{X_n}(o_j)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_m	$e_{X_1}(o_m)$	$\cdots$	$e_{X_i}(o_m)$	$\cdots$	$e_{X_n}(o_m)$

For each X_i specify:

- context w_i,
- a number s_i of nodes,
- set $D_i = \{y_{ik} \in w_i \mid k = 1, \ldots, s_i\}$ of nodes,
- fuzzy partition $\{\mathrm{Fn}_{w_i}(y_{ik}) \underset{\sim}{\subset} w_i \mid y_{ik} \in D_i\}$.

Data (again):

	X_1	$\cdots$	X_i	$\cdots$	X_n
o_1	$e_{X_1}(o_1)$	$\cdots$	$e_{X_i}(o_1)$	$\cdots$	$e_{X_n}(o_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_j	$e_{X_1}(o_j)$	$\cdots$	$e_{X_i}(o_j)$	$\cdots$	$e_{X_n}(o_j)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_m	$e_{X_1}(o_m)$	$\cdots$	$e_{X_i}(o_m)$	$\cdots$	$e_{X_n}(o_m)$

For each X_i specify:

- context w_i,
- a number s_i of nodes,
- set $D_i = \{y_{ik} \in w_i \mid k = 1, \ldots, s_i\}$ of nodes,
- fuzzy partition $\{\mathrm{Fn}_{w_i}(y_{ik}) \underset{\sim}{\subset} w_i \mid y_{ik} \in D_i\}$.

Antecedent: $X_1, \ldots, X_p$, Consequent: Z.

$D = D_1 \times \cdots \times D_p$ — set of all p-tuples of nodes.

$\bar{y} = \langle y_{1k_1}, \ldots, y_{pk_p} \rangle \in D$ — elements (vectors of nodes)

Form of associations:

$$(X_1 \text{ is } \mathrm{Fn}(y_{1k_1})) \text{ AND} \cdots \text{AND} (X_p \text{ is } \mathrm{Fn}(y_{pk_p})) \overset{F}{\sim}_{r,\gamma} (\texttt{average } Z \text{ is } \mathcal{B}_{\bar{y}}),$$

Antecedent: Multidimensional fuzzy number

$$A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j)) = \mathrm{Fn}(y_{1k_1}, e_1(o_j)) \cdots \mathrm{Fn}(y_{pk_p}, e_p(o_j))$$

Form of associations

Antecedent: $X_1, \ldots, X_p$, Consequent: Z.
$D = D_1 \times \cdots \times D_p$ — set of all p-tuples of nodes.
$\bar{y} = \langle y_{1k_1}, \ldots, y_{pk_p} \rangle \in D$ — elements (vectors of nodes)

Form of associations:

$$(X_1 \text{ is } \mathrm{Fn}(y_{1k_1}))\ \mathrm{AND} \cdots \mathrm{AND} (X_p \text{ is } \mathrm{Fn}(y_{pk_p}))$$
$$\overset{F}{\sim}_{r,\gamma}\ (\texttt{average } Z \text{ is } \mathcal{B}_{\bar{y}}),$$

Antecedent: Multidimensional fuzzy number

$$A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j)) = \mathrm{Fn}(y_{1k_1}, e_1(o_j)) \cdots \mathrm{Fn}(y_{pk_p}, e_p(o_j))$$

Form of associations

Antecedent: $X_1, \ldots, X_p$, Consequent: Z.

$D = D_1 \times \cdots \times D_p$ — set of all p-tuples of nodes.

$\bar{y} = \langle y_{1k_1}, \ldots, y_{pk_p} \rangle \in D$ — elements (vectors of nodes)

Form of associations:

$$(X_1 \text{ is } \mathrm{Fn}(y_{1k_1}))\,\text{AND} \cdots \text{AND}(X_p \text{ is } \mathrm{Fn}(y_{pk_p}))$$
$$\overset{F}{\sim}_{r,\gamma} (\texttt{average } Z \text{ is } \mathcal{B}_{\bar{y}}),$$

Antecedent: Multidimensional fuzzy number

$$A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j)) = \mathrm{Fn}(y_{1k_1}, e_1(o_j)) \cdots \mathrm{Fn}(y_{pk_p}, e_p(o_j))$$

Consequent

Value of consequent Z:

$$F_{\bar{y}} = \frac{\sum_{j=1}^{m} A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j)) \cdot e_Z(o_j)}{\sum_{j=1}^{m} A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j))}.$$

Perception of $F_{\bar{y}}$ in the context w_Z:

$$\mathcal{B}_{\bar{y}} = \mathrm{Suit}(F_{\bar{y}}, w_Z).$$

Result:

`average` Z is $\mathcal{B}_{\bar{y}}$.

Consequent

Value of consequent Z:

$$F_{\bar{y}} = \frac{\sum_{j=1}^{m} A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j)) \cdot e_Z(o_j)}{\sum_{j=1}^{m} A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j))}.$$

Perception of $F_{\bar{y}}$ in the context w_Z:

$$\mathcal{B}_{\bar{y}} = \mathrm{Suit}(F_{\bar{y}}, w_Z).$$

Result:

$$\texttt{average } Z \text{ is } \mathcal{B}_{\bar{y}}.$$

Consequent

Value of consequent Z:

$$F_{\bar{y}} = \frac{\sum_{j=1}^{m} A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j)) \cdot e_Z(o_j)}{\sum_{j=1}^{m} A_{\bar{y}}(\bar{e}_{\mathcal{Y}}(o_j))}.$$

Perception of $F_{\bar{y}}$ in the context w_Z:

$$\mathcal{B}_{\bar{y}} = \mathrm{Suit}(F_{\bar{y}}, w_Z).$$

Result:

$$\texttt{average}\ Z \text{ is } \mathcal{B}_{\bar{y}}.$$

Purpose of LFLC2000 software

LFLC2000 (Linguistic Fuzzy Logic Controller) is an universal software system dedicated primarily for designing and testing of linguistic descriptions, i.e. systems of fuzzy IF-THEN rules.

Originated by Vilém Novák in 1990's.
Developed in IRAFM, University of Ostrava.

LFLC offers unique methodology based on theoretical achievements from IRAFM members.

Data for testing

NO_2 data, Oslo, Norway

- X_1 — logarithm of the number of cars per hour,
- X_2 — temperature 2 meter above ground (*degree C*),
- X_3 — wind speed (*meters/second*),
- X_4 — the temperature difference between 25 and 2 meters above ground (*degree C*),
- X_5 — wind direction (*degrees between 0 and 360*),
- Z — The response variable — hourly values of the logarithm of the concentration of NO_2

NO_2 linguistic associations

A1. $(X_1 \text{ is } MLMe) \text{ AND} (X_2 \text{ is } VeSm) \text{ AND} (X_3 \text{ is } MLSm) \sqsubset^{0.5}_{0.01} (Z \text{ is } MLMe)$

A2. $(X_1 \text{ is } VeBi) \text{ AND} (X_2 \text{ is } MLSm) \text{ AND} (X_3 \text{ is } MLSm) \sqsubset^{0.5}_{0.01} (Z \text{ is } MLBi)$

A3. $(X_1 \text{ is } Bi) \text{ AND} (X_2 \text{ is } \text{-}MLSm) \text{ AND} (X_3 \text{ is } MLMe) \sqsubset^{0.5}_{0.01} (Z \text{ is } MLBi)$

A4. $(X_1 \text{ is } Bi) \text{ AND} (X_2 \text{ is } MLSm) \text{ AND} (X_3 \text{ is } MLMe) \sqsubset^{0.5}_{0.01} (Z \text{ is } MLBi)$

Reduction of A3 and A4 to

$$(X_1 \text{ is } Bi \text{ AND} (X_2 \text{ is } (MLSm\ OR\ \text{-}MLSm)) \text{ AND} (X_3 \text{ is } MLMe) \sqsubset^{0.5}_{0.01} (Z \text{ is } MLBi)$$

NO_2 FT associations

B1. $(X_1 \text{ is } \mathrm{Fn}_{w_1}(7.65))\,\text{AND}(X_2 \text{ is } \mathrm{Fn}_{w_2}(1.25))$
$\text{AND}(X_3 \text{ is } \mathrm{Fn}_{w_3}(0.3)) \overset{F}{\sim}_{0.1,0.2} (\texttt{average}\ \ Z \text{ is } Bi)$

B2. $(X_1 \text{ is } \mathrm{Fn}_{w_1}(7.65))\,\text{AND}(X_2 \text{ is } \mathrm{Fn}_{w_2}(-5.37))$
$\text{AND}(X_3 \text{ is } \mathrm{Fn}_{w_3}(3.5)) \overset{F}{\sim}_{0.1,0.2} (\texttt{average}\ \ Z \text{ is } Me)$

B3. $(X_1 \text{ is } \mathrm{Fn}_{w_1}(7.65))\,\text{AND}(X_2 \text{ is } \mathrm{Fn}_{w_2}(7.87))$
$\text{AND}(X_3 \text{ is } \mathrm{Fn}_{w_3}(5.1)) \overset{F}{\sim}_{0.1,0.2} (\texttt{average}\ \ Z \text{ is } QRBi)$

Conclusions & future work

Future work

- Generalized quantifiers
- Automatic selection of variables
- Automatic selection of contexts
- Theoretical analysis of reduction rules
- Computational complexity of algorithms involved
- Testing on larger data files