Using FCA and Concept Explications for Finding an Appropriate Concept

Raslan 2021

Marek Menšík[1], **Adam Albert**[1], Tomáš Michalovský[2]

VŠB – Technická univerzita Ostrava[1],
Silesian University in Opava[2]

December 10, 2021

Explication

- Explication is the process of refinement of a vague or inaccurate expression into an adequately accurate one.
- Symbolic methods of Supervised machine learning + Transparent Intensional Logic are used to obtain explications of atomic concepts

$$'Dog \approx_{exp} \lambda w \lambda t \, \lambda x \, [['Domesticated \;\; 'Carnivore]_{wt} \, x]$$

Types: *Domesticated*$/((\iota)_{wt}(\iota)_{wt})$; *Dog*, *Carnivore*$/(\iota)_{wt}$; $x \rightarrow \iota$

An Example of Explication

$$e_1 = [{}^{\prime}Typp\ \lambda w \lambda t\ \lambda x[[{}^{\prime}\leq [{}^{\prime}Weight_{wt}\ x]\ {}^{\prime}11]\ \wedge\ [{}^{\prime}\geq [{}^{\prime}Weight_{wt}\ x]\ {}^{\prime}1.2]][{}^{\prime}Wild\ {}^{\prime}Cat]]\ \wedge\ [{}^{\prime}Req\ {}^{\prime}Mammal\ [{}^{\prime}Wild\ {}^{\prime}Cat]]\ \wedge\ [{}^{\prime}Req\ {}^{\prime}Hasfur\ [{}^{\prime}Wild\ {}^{\prime}Cat]] \wedge [{}^{\prime}Typp\ \lambda w \lambda t\ \lambda x[[{}^{\prime}\leq [[{}^{\prime}Average\ {}^{\prime}BodyLength]\ x]\ {}^{\prime}80]\ \wedge\ [{}^{\prime}\geq [[{}^{\prime}Average\ {}^{\prime}BodyLength]\ x]\ {}^{\prime}47]][{}^{\prime}Wild\ {}^{\prime}Cat]] \wedge [{}^{\prime}Typp\ \lambda w \lambda t\ \lambda x[[{}^{\prime}= [[{}^{\prime}Average\ {}^{\prime}SkullSize]\ x]\ {}^{\prime}41.25]][{}^{\prime}Wild\ {}^{\prime}Cat]]\ \wedge\ [{}^{\prime}Typp\ \lambda w \lambda t\ \lambda x[[{}^{\prime}= [[{}^{\prime}Average\ {}^{\prime}Height]\ x]\ {}^{\prime}37,6]][{}^{\prime}Wild\ {}^{\prime}Cat]]$$

Supervised machine learning

- Examples described by (input/output) attributes provided by a supervisor.
- Learner is building his hypothesis based on seen examples.
- Prediction of output attribute values based on the values of input attributes.
- Training/test examples.

Algorithm framework

- Machine learning algorithm can be described by:
  - Task goal
    - The goal is to discover the best refinement of the learned simple concept.
  - Training data
    - The learner is working with a set of positive and negative examples.
    - Positive examples contain concept requisites.
    - Negative examples specify properties which are not in the essence of a concept.
  - Data representation
    - Transparent Intensional Logic (TIL).
  - Knowledge modifying module $\rightarrow$ three heuristic methods:
    - Specialization.
    - Generalization.
    - Refinement.

Knowledge modifying module

- Negative examples $\rightarrow$ *Specialization* inserts negated constituents.
- Positive examples:
  - $\rightarrow$ *Refinement* inserts new constituents into the molecular construction learned so far.
  - $\rightarrow$ *Generalization* adjusts the constituents.

Specialization

If the negative example has a constituent that the model does not have.

- Hypothesis: *Cat is a domesticated feline predatory mammal.*
- Negative example: *Dog is a domesticated predatory mammal that barks.*

$$\lambda w \lambda t \lambda x \left[\left[\left[{}'Feline\ [{}'Predatory\ {}'Mammal] \right]_{wt}\ x \right] \wedge \left[{}'Domesticated_{wt}\ x \right] \wedge \neg [{}'Bark_{wt}\ x] \wedge \neg \left[{}'Dog_{wt}\ x \right] \right]$$

Refinement

If the positive example contains a constituent that the model does not have.

- Hypothesis: *Cat is a domesticated feline predatory mammal and does not bark and is not a dog.*
- Positive example: *Cat has fur.*

$$\lambda w \lambda t \lambda x \left[\left[\left['Feline \; ['Predatory \; 'Mammal] \right]_{wt} \; x \right] \wedge \left['Domesticated_{wt} \; x \right] \wedge \left['Hasfur_{wt} \; x \right] \wedge \neg ['Bark_{wt} \; x] \wedge \neg \left['Dog_{wt} \; x \right] \right]$$

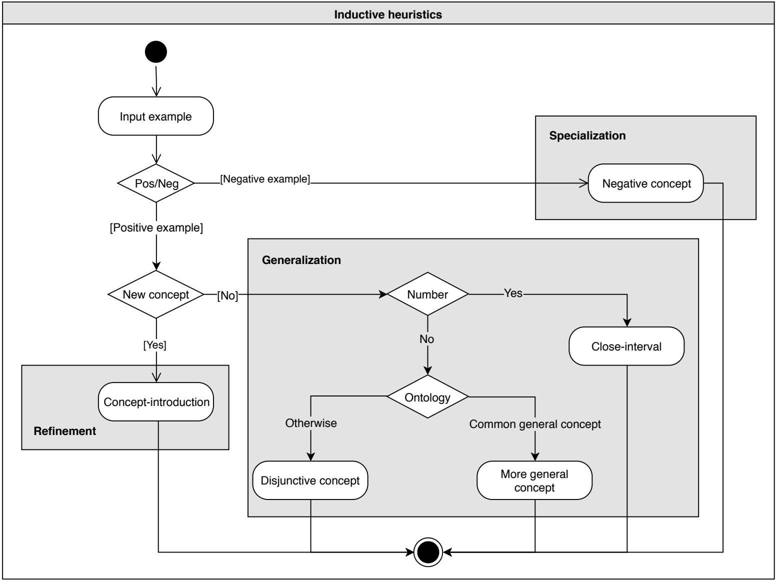
Generalization

If the constituents of model and example differ and constituents do not have the most specific general concept.
Generalization is achieved by unifying these classes.

- Hypothesis: *Cat is a domesticated feline predatory mammal which is wild or domesticated and does not bark and is not a dog and has a fur.*
- Positive example: *Cat is a wild feline predatory mammal.*

$$\lambda w \lambda t \lambda x \left[\left[\left[{}'Feline \; [{}'Predatory \; {}'Mammal] \right]_{wt} \; x \right] \wedge \left[[{}'Domesticated_{wt} \; x] \vee [{}'Wild_{wt} \; x] \right] \wedge \left[{}'Has - fur_{wt} \; x \right] \wedge \neg [{}'Bark_{wt} \; x] \wedge \neg \left[{}'Dog_{wt} \; x \right] \right]$$

Algorithm diagram



Formal Conceptual Analysis - FCA

- FCA has been applied in many disciplines such as software engineering, machine learning, knowledge discovery and ontology construction.
- Informally, FCA studies how objects can be hierarchically grouped together with their mutual common attributes.
- FCA is utilized to obtain all formal concepts and create conceptual lattice over explications.
- Formal context: (G, M, I)
- The set of formal concepts:

$$\beta(G, M, I) = \{(O, A) | O \subseteq G, A \subseteq M, A^{\downarrow} = O, O^{\uparrow} = A\}$$

  - $O^{\uparrow} = \{a | \forall o \in O, (o, a) \in I\}$
  - $A^{\downarrow} = \{o | \forall a \in A, (o, a) \in I\}$

Aspirant Ordering

Definition

(1) **Concept aspirants** of the set of attributes **a** in $\beta(G, M, I)$ is a set $CA(a) = \bigcup_{i=1}^{n} O_i^a$, where O^a is extent of a concept $(O, A) \neq (G, B), a \subseteq A, B \subseteq M$. Namely, **concept aspirants** of the set of attributes **a** is a union of all formal concept extents where **a** is a subset of a particular formal concepts' intents.

Definition

(2) Let $CA(a)$ be a set of concept aspirants of a set of attributes **a**, let $\delta(a)$ be a set of concepts (O, A) where $a \subseteq A$, i.e.: $\delta(a) = \{(O^a, (O^a)^\uparrow) | (O^a, (O^a)^\uparrow) \neq (G, B), B \subseteq M, (O^a, (O^a)^\uparrow) \in \beta(G, M, I)\}$. Then **x** $\sqsubseteq$ **y** is in relation of **aspirant ordering** iff $max(|(O^y)^\uparrow|) \leq max(|(O^x)^\uparrow|), x, y, \in CA(a), (O^x, (O^x)^\uparrow), (O^y, (O^y)^\uparrow) \in \delta(a)$.

Most appropriate concepts

Definition

(3) Let $(CA(a), \sqsubseteq)$ be an ordered set according to the definition 2, then the maximal elements are the **most appropriate concepts**.

Table: Formal context

	a_0	a_1	a_2	a_3
o_0	1	1	0	0
o_1	0	1	1	0
o_2	0	1	1	1

The set of all *formal concepts*
(G,M,I) = { $\mathsf{C}_0, C_1, C_2, C_3, C_4\}$, where

$$C_0 = (\{o_0, o_1, o_2\}, \{a_1\}) \quad C_1 = (\{o_0\}, \{a_0, a_1\})$$
$$C_2 = (\{o_1, o_2\}, \{a_1, a_2\}) \quad C_3 = (\{o_2\}, \{a_1, a_2, a_3\})$$
$$C_4 = (\emptyset, \{a_0, a_1, a_2, a_3\})$$

Frame Title

Find the set of *concept aspirants* for attributes $a = \{a_2\}$

1. Find set $\delta(a)$:
   $\delta(a) = \{(\{o_1, o_2\}, \{a_1, a_2\}), (\{o_2\}, \{a_1, a_2, a_3\}), (\emptyset, \{a_0, a_1, a_2, a_3\})\}$
2. Create the union of all extents found in step 1 : $CA(\{a_2\}) = \{o_1, o_2\}$
3. For all $x \in CA(\{a_2\})$ calculate max of $|(O^x)^{\uparrow}|$, where $((O^x), (O^x)^{\uparrow}) \in \delta(a) \rightarrow max(|\{o_1, o_2\}^{\uparrow}|) = 2$, $max(\{|\{o_1, o_2\}^{\uparrow}|, |o_2{}^{\uparrow}|\}) = 3$
4. Order $CA(\{a_2\})$ by definition 2 $\rightarrow o_2 \sqsubseteq o_1$

Table: Aspirants' ordering

Explication	Intent	DF
o_1	$\{a_1, a_2\}$	$\{a_1\}$
o_2	$\{a_1, a_2, a_3\}$	$\{a_1, a_3\}$

DF represents the difference from the selected set of attributes.

Case study

Table: Explications and attributes

Explication (O)	Attributes (A)
Jungle cat (JC)	$\{a_1, a_2, a_3, a_4, a_5\}$
Sand cat (SC)	$\{a_1, a_2, a_4, a_6, a_7\}$
House cat (HC)	$\{a_1, a_2, a_8, a_9\}$
Lynx (Ly)	$\{a_1, a_2, a_{10}, a_{11}, a_{12}\}$
Lion (Li)	$\{a_1, a_2, a_{13}, a_{14}, a_{17}, a_{18}\}$
Tiger (Ti)	$\{a_1, a_2, a_{15}, a_{16}, a_{17}\}$

Table: Table of all formal concepts

C	Extent	Intent
C_1	O	$\{a_1, a_2\}$
C_2	$\{JC, SC\}$	$\{a_1, a_2, a_4\}$
C_3	$\{Li, Ti\}$	$\{a_1, a_2, a_{17}\}$
C_4	$\{HC\}$	$\{a_1, a_2, a_8, a_9\}$
C_5	$\{JC\}$	$\{a_1, a_2, a_3, a_4, a_5\}$
C_6	$\{SC\}$	$\{a_1, a_2, a_4, a_6, a_7\}$
C_7	$\{Ly\}$	$\{a_1, a_2, a_{10}, a_{11}, a_{12}\}$
C_8	$\{Ti\}$	$\{a_1, a_2, a_{15}, a_{16}, a_{17}\}$
C_9	$\{Li\}$	$\{a_1, a_2, a_{13}, a_{14}, a_{17}, a_{18}\}$
C_{10}	$\emptyset$	A

Case study

Table: Table of all formal concepts

a_1	$'Mammal$
a_2	$'Hasfur$
a_3	$\lambda w \lambda t\ \lambda x[[' \leq ['Bdlgth_{wt}\ x]\ '112] \wedge\ [' \geq ['Bdlght_{wt}\ x]\ '55]]$
a_4	$\lambda w \lambda t\ \lambda x\, [' =_p ['Furcolor_{wt}\ x]\ \ 'Brown]$
a_5	$\lambda w \lambda t\ \lambda x\, [' = [['Avg\ 'Height]_{wt}\ \ x]\ \ '36.5]$
a_6	$\lambda w \lambda t\ \lambda x[[' \leq ['Bdlgth_{wt}\ x]\ '57] \wedge\ [' \geq ['Bdlght_{wt}\ x]\ '39]]$
a_7	$\lambda w \lambda t\ \lambda x\, [' = [['Avg\ 'Height]_{wt}\ \ x]\ \ '27]$
a_8	$'Domesticated$
a_9	$\lambda w \lambda t\ \lambda x\, [' = [['Avg\ 'Height]_{wt}\ \ x]\ \ '30]$
a_{10}	$\lambda w \lambda t\ \lambda x\, [' \leq [['Avg\ 'Bdlgth]_{wt}\ \ x]\ \ '148]$
a_{11}	$\lambda w \lambda t\ \lambda x\, [' = [['Avg\ 'Height]_{wt}\ \ x]\ \ '75]$
a_{12}	$['Biggest\ \ ['EU\ \ ['Feline\ 'Predator]]]$
a_{13}	$'Hasmane$
a_{14}	$\lambda w \lambda t\ \lambda x[[' \leq ['Bdlgth_{wt}\ x]\ '250] \wedge\ [' \geq ['Bdlght_{wt}\ x]\ '170]]$
a_{15}	$['Apex\ 'Predator]$
a_{16}	$\lambda w \lambda t\ \lambda x\, [' = [['Avg\ 'Height]_{wt}\ \ x]\ \ '117]$
a_{17}	$'Pantherinae$
a_{18}	$['Significant\ 'SexDimorph]$

Case study

- Chosen attribute:

$$CA(\{a_{17}\}) = \{Li, Ti\}$$

- The set $CA(\{a_{17}\})$ is ordered according to definition 2. The final ordering is as follows:

$$Li \sqsubseteq Ti$$

- According to definition 3 the entity *Ti* is is a maximal one, and thus the concept of *'being a Tiger'* is presented to the user as the most appropriate one.

Conclusion

- Symbolic method of supervised machine learning.
- Formal Conceptual Analysis over explications created by the machine learning algorithm.
- Ordering of concept aspirants using attributes properties and attributes' values known by user.
- Finding an appropriate concept.

Thank you for your attention

Marek Menšík[1], **Adam Albert**[1],Tomáš Michalovský[2]

VŠB – Technická univerzita Ostrava[1],
Silesian University in Opava[2]

December 10, 2021