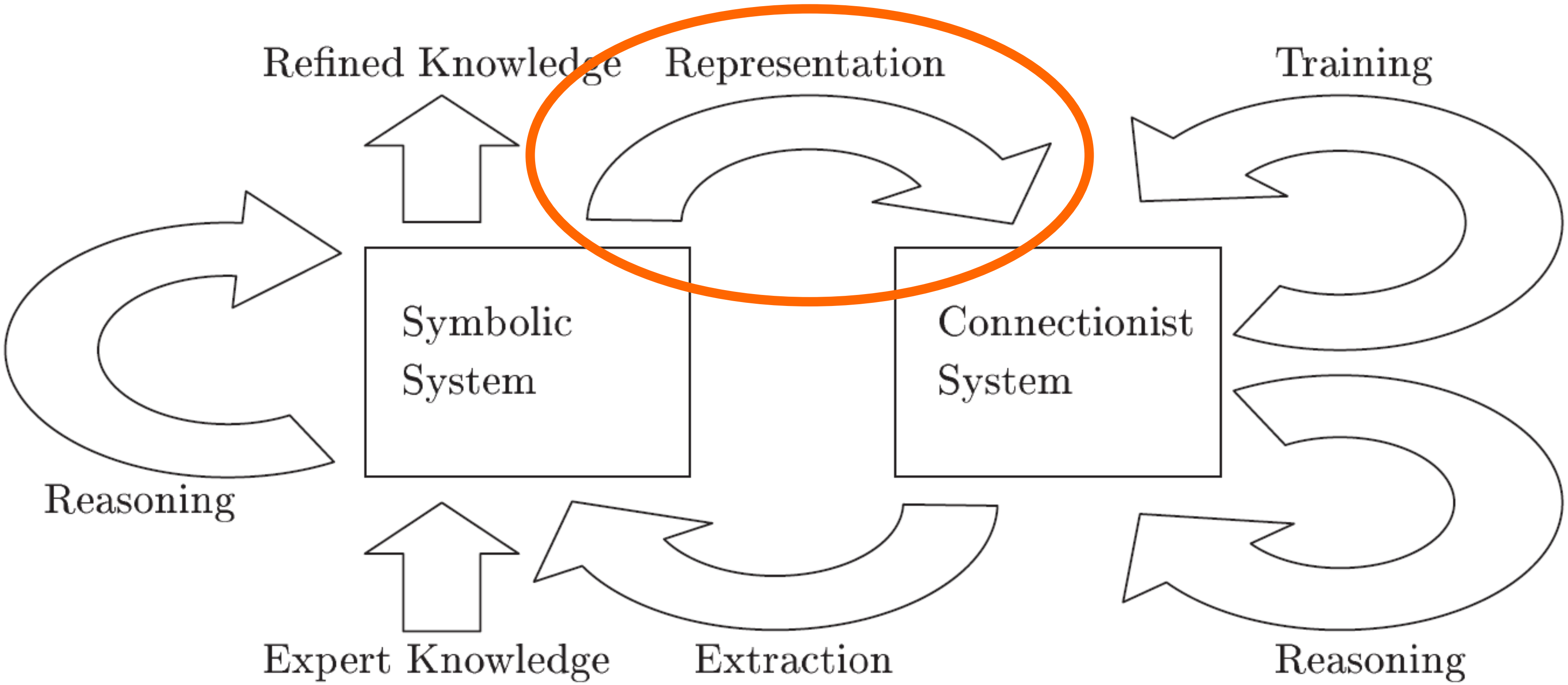
AIFB

Encoding Closure Operators into Neural Networks

Sebastian Rudolph
University of Karlsruhe
Institute AIFB
rudolph@aifb.uni-karlsruhe.de

AIFB

Placement in the big picture



Outline

- Closure Operators
- Formal Concept Analysis – Basics
- Two Ways for Encoding
- Example & Relation to Alternative Approach

AIFB

What is a closure operator?

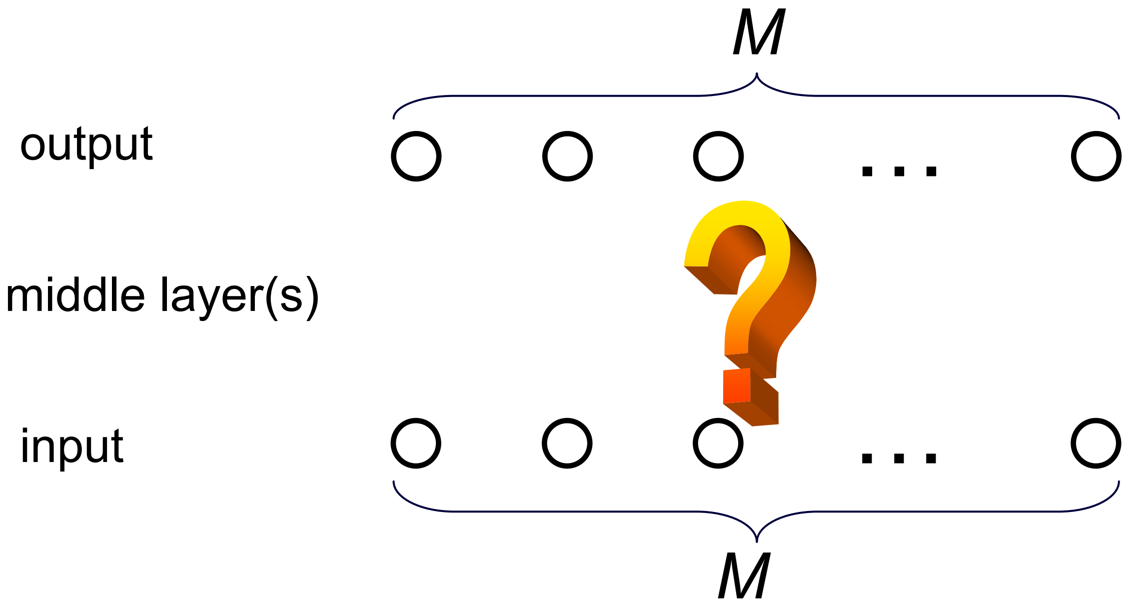
- Given: base set M
- Function φ: $2^M \rightarrow 2^M$ called closure operator if
  - extensive, i.e. $A \subseteq \varphi(A)$
  - monotone, i.e. $A \subseteq B$ implies $\varphi(A) \subseteq \varphi(B)$
  - idempotent, i.e. $\varphi(A) = \varphi(\varphi(A))$
- Example:
  - M set of all FOL formulae for a given signature
  - set of consequences of a given set of FOL formulae

(…in fact, every monotonistic logic is fine)

AIFB

Closure operators in neural networks

- Question: How to realize a closure operator in a neural network?
- input and output layer clear, but what about the rest?



AIFB

FCA Basics – Formal Contexts

formal context: $\mathbb{K}=(G,M,I)$

- set G objects
- set M attributes
- $I \subseteq G \times M$

gIm interpreted as:
„object g has attribute m"

Example:

	c0	*c1*	*c2*	*g2*	*pr*	*ev*	*od*
0	×					×	
1		×					×
2			×		×	×	
3				×	×		×
4				×		×	
5				×	×		×
6				×		×	
7				×	×		×
8				×		×	
9				×			×
...							

ev	even
od	odd
pr	prime
c0	equals zero
c1	equals one
c2	equals two
g2	greater than two

FCA Basics – Derivation Operators

Given $\mathbb{K}=(G,M,I)$, define
for $A \subseteq G$ and $B \subseteq M$

- $A^I := \{m \mid gIm \text{ for all } g \in A\}$
  (all attributes common to all objects of A)

- $B^I := \{g \mid gIm \text{ for all } m \in B\}$
  (all objects having every attribute of B)

Example: $\{2,4,6\}^I = \{ev\}$

M

G

	e0	*e1*	*e2*	*g2*	*pr*	*ev*	*od*
0	×					×	
1		×					×
2			×		×	×	
3				×	×		×
4				×		×	
5				×	×		×
6				×		×	
7				×	×		×
8				×		×	
9				×			×
...							

ev	even
od	odd
pr	prime
e0	equals zero
e1	equals one
e2	equals two
g2	greater than two

FCA Basics – Formal Concepts

formal concept: pair (A,B)
with $A^I=B$ and $B^I=A$

Formal concepts of K
- can be ordered via
  $(A,B)\leq(C,D)$ iff $A\subseteq C$
- constitute a complete lattice.

Example:

$(\{2n+3 \mid n\in \mathbb{N}\}, \{g2,od\})$

M

G

	e0	e1	e2	g2	pr	ev	od
0	×					×	
1		×					×
2			×		×	×	
3				×	×		×
4				×		×	
5				×	×		×
6				×		×	
7				×	×		×
8				×		×	
9				×			×
...							

ev	even
od	odd
pr	prime
e0	equals zero
e1	equals one
e2	equals two
g2	greater than two

FCA Basics – Formal Concepts

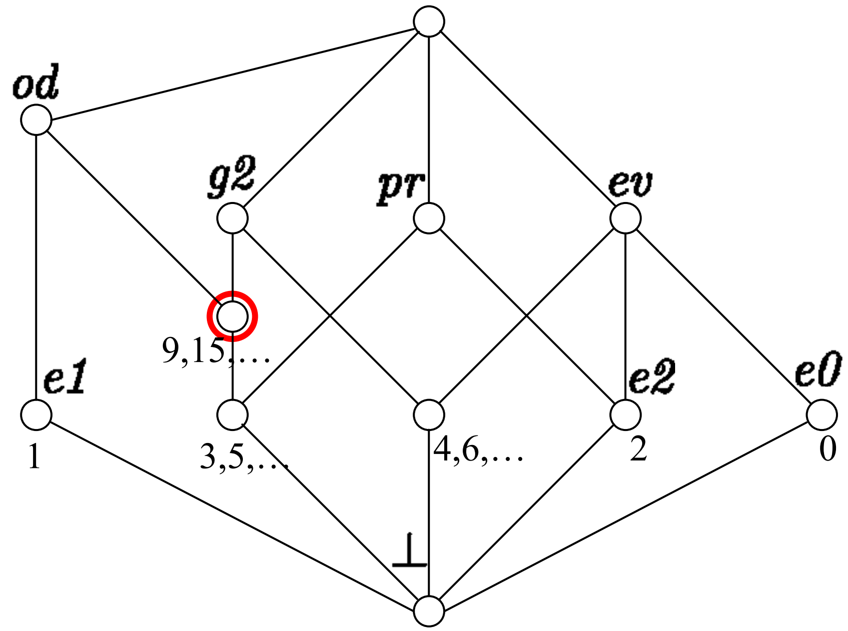
formal concept: pair (A,B)
with $A^I=B$ and $B^I=A$

Formal concepts of K

- can be ordered via

  $(A,B)\leq(C,D)$ iff $A\subseteq C$

- constitute a complete lattice.



Example:

$(\{2n+3 \mid n\in \mathbb{N}\}, \{g2,od\})$

ev	even
od	odd
pr	prime
e0	equals zero
e1	equals one
e2	equals two
g2	greater than two

FCA Basics – Attribute Implications

Let $A,B \subseteq M$.
Implication $A \rightarrow B$ holds in $\mathbb{K}=(G,M,I)$, if for every $g \in G$
$A \subseteq g^I$ implies $B \subseteq g^I$

$g^I := \{m \in M \mid gIm\}$
(= all attributes of object g)

Example:

M (attributes), G (objects)

	e0	e1	e2	g2	pr	ev	od
0	×					×	
1		×					×
2			×		×	×	
3				×	×		×
4				×		×	
5				×	×		×
6				×		×	
7				×	×		×
8				×		×	
9				×			×
...							

ev	even
od	odd
pr	prime
e0	equals zero
e1	equals one
e2	equals two
g2	greater than two

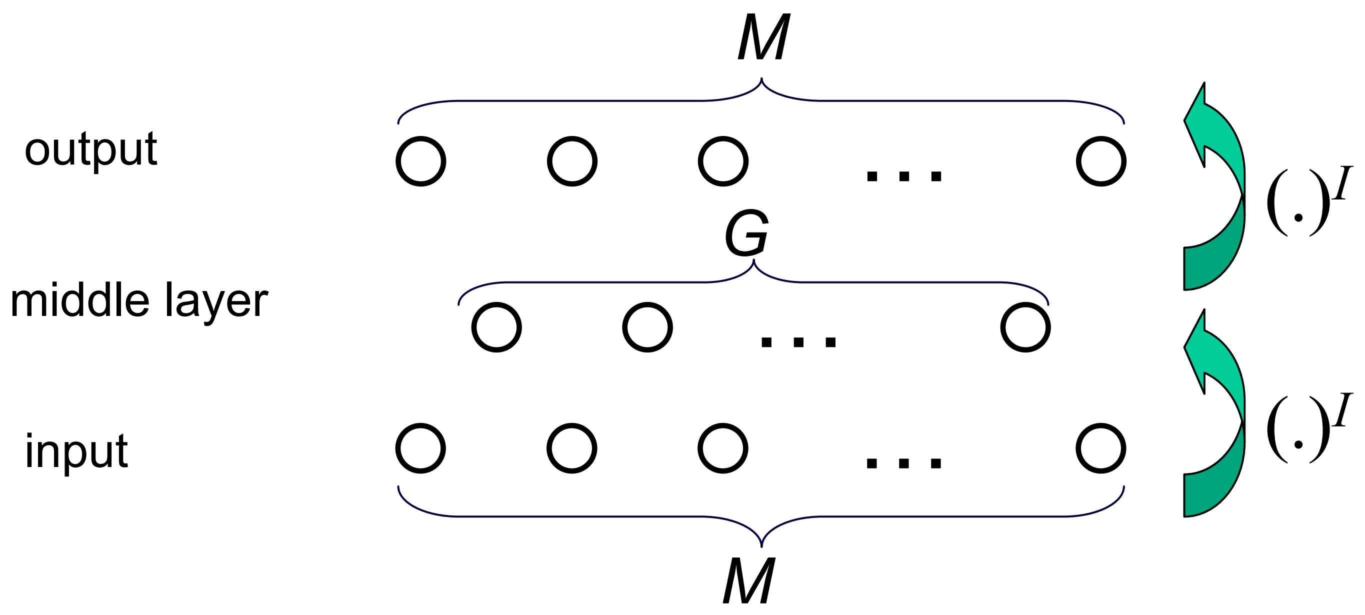
Some implications valid in the example:

$$\{g2,pr\} \rightarrow \{od\}$$
$$\{e0,e1\} \rightarrow M$$

AIFB

Now how can FCA help?

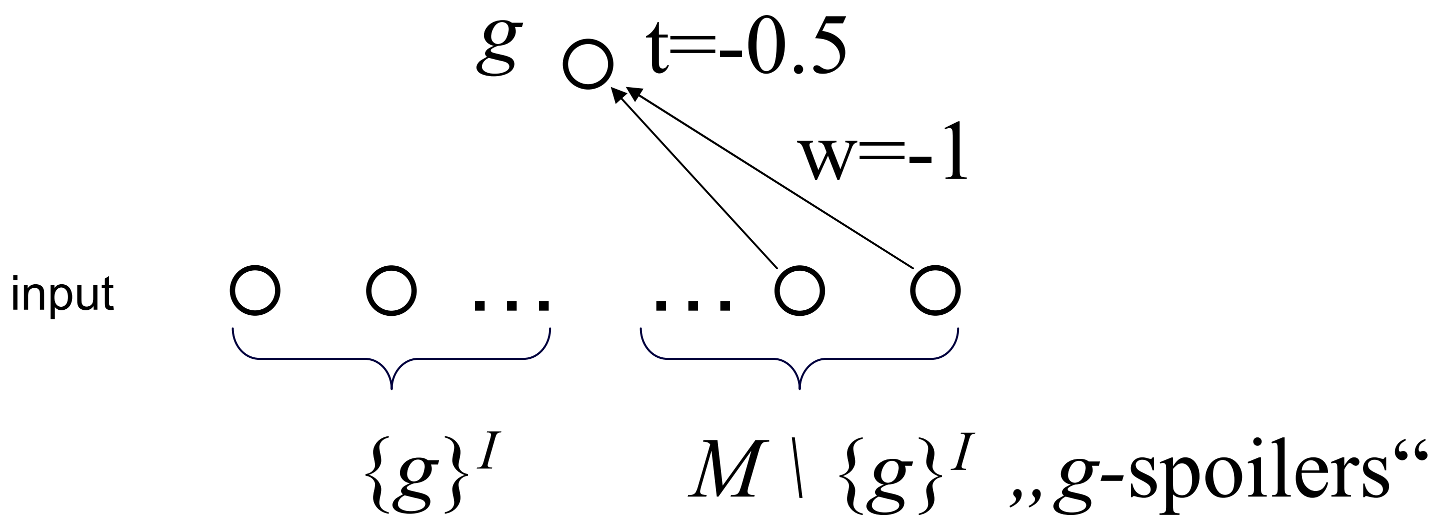
- In a formal context (G,M,I), the function
  $(.)^{II} : 2^M \to 2^M ,\ A \mapsto A^{II}$
  is a closure operator on the attribute set.
- Idea:



AIFB

What about links, weights and thresholds?

- middle layer neuron associated to object *g* activated exactly if $g \in A^I$
- this is equivalent to $A \subseteq \{g\}^I$

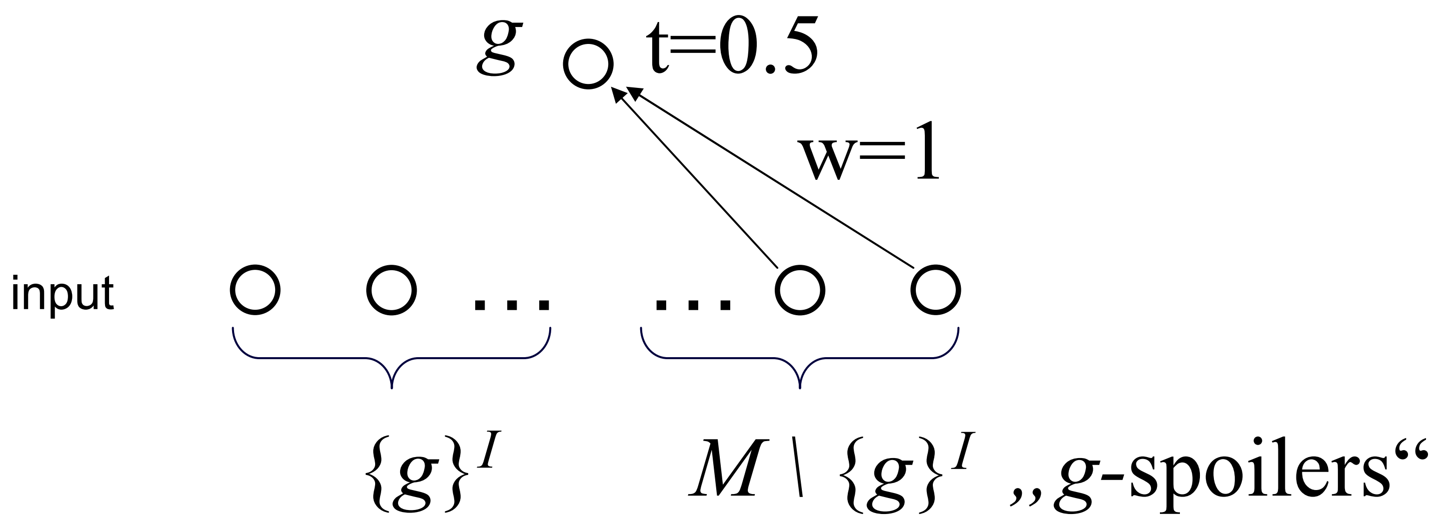


AIFB

- twofold linking in this way yields desired neural network which calculates $(.)^{II}$
- negative weights necessary because $(.)^{I}$ is antitone (i.e. $A \subseteq B$ implies $B^I \subseteq A^I$)
- Can we do better?
- Yes! For input *A*, activate those middle layer neurons *not* belonging to A^I (i.e., activate $M \setminus A^I$). This makes mappings monotone and hence allows for positive weights and thresholds.

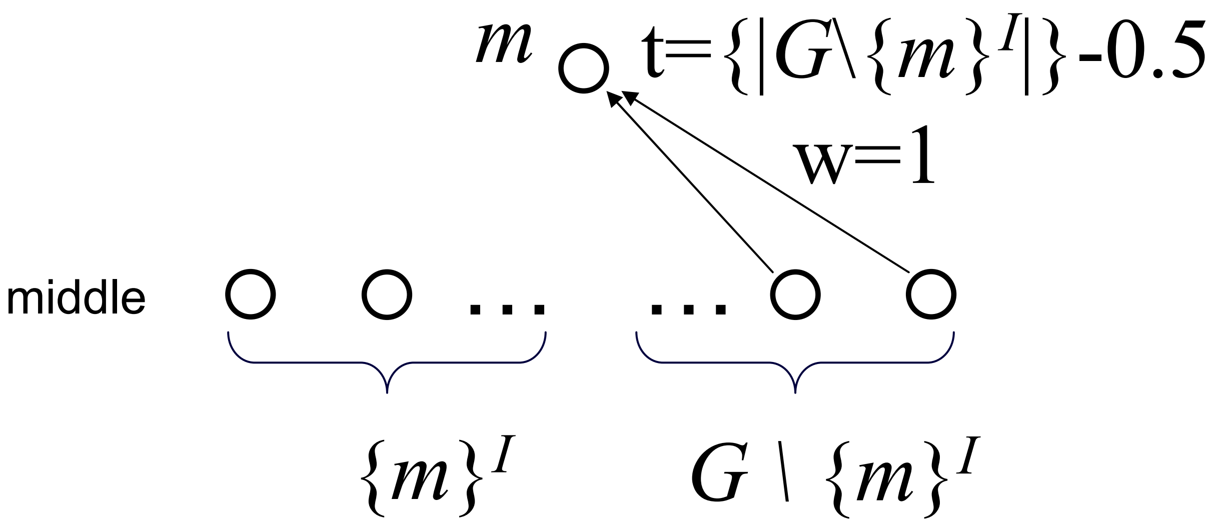
Middle layer revisited…

- middle layer neuron associated to object *g* activated exactly if not $g \in A^I$



Output layer revisited…

- with B set of activated middle layer neurons, output layer neuron associated to attribute m activated exactly if $m \in (G \setminus B)^I$
- equivalent to $(G \setminus B) \subseteq \{m\}^I$
- equivalent to $(G \setminus \{m\}^I) \subseteq B$



Example

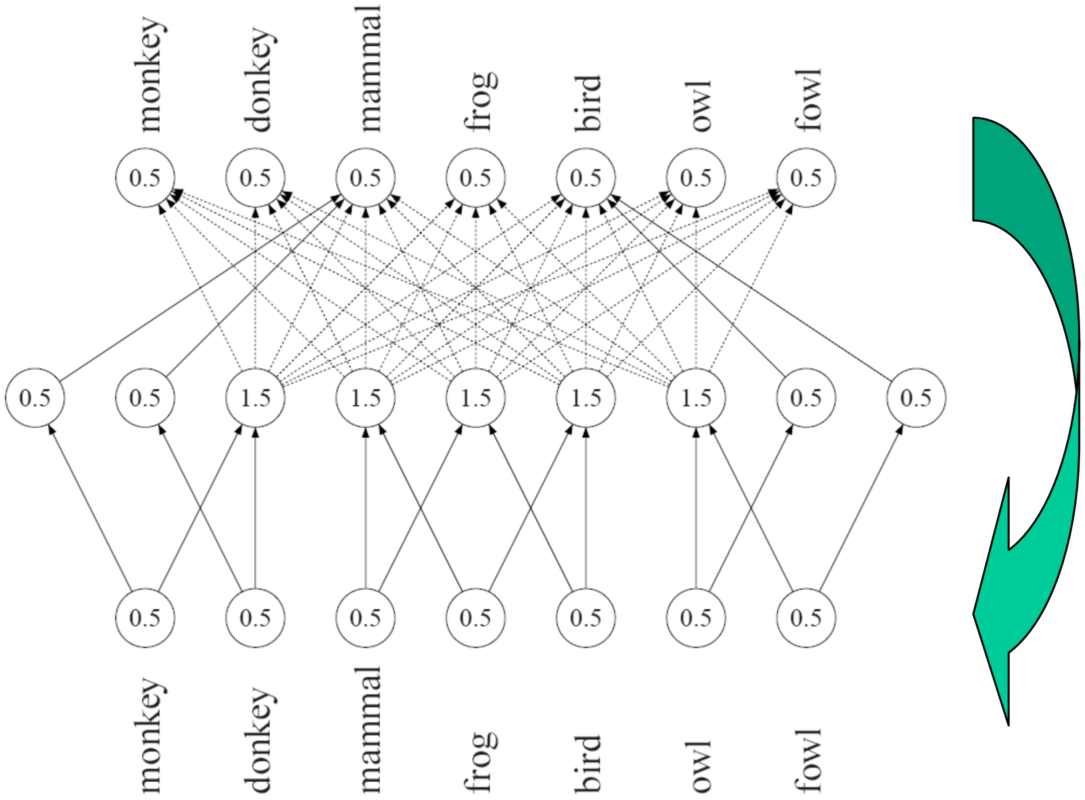
- consider propositional logic program:

$$
\begin{aligned}
&\text{monkey} \multimap \text{mammal}\\
&\text{donkey} \multimap \text{mammal}\\
&\text{owl} \multimap \text{bird}\\
&\text{fowl} \multimap \text{bird}\\
&\text{monkey, donkey} \multimap \bot\\
&\text{owl, fowl} \multimap \bot\\
&\text{mammal, bird} \multimap \bot\\
&\text{mammal, frog} \multimap \bot\\
&\text{bird, frog} \multimap \bot
\end{aligned}
$$

AIFB

Example

- as solved by [Hölldobler & Kalinke] (recurrent version):

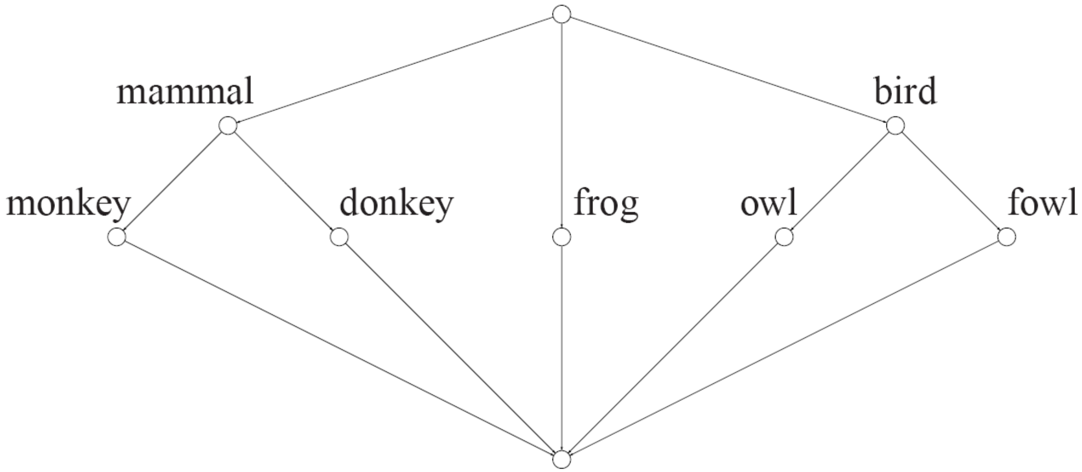


AIFB

Example

- as solved via FCA:

$$
\begin{array}{ll}
\text{monkey} & \to \text{mammal} \\
\text{donkey} & \to \text{mammal} \\
\text{owl} & \to \text{bird} \\
\text{fowl} & \to \text{bird} \\
\text{monkey, donkey} & \to \bot \\
\text{owl, fowl} & \to \bot \\
\text{mammal, bird} & \to \bot \\
\text{mammal, frog} & \to \bot \\
\text{bird, frog} & \to \bot
\end{array}
$$



AIFB

Example

- as solved via FCA:

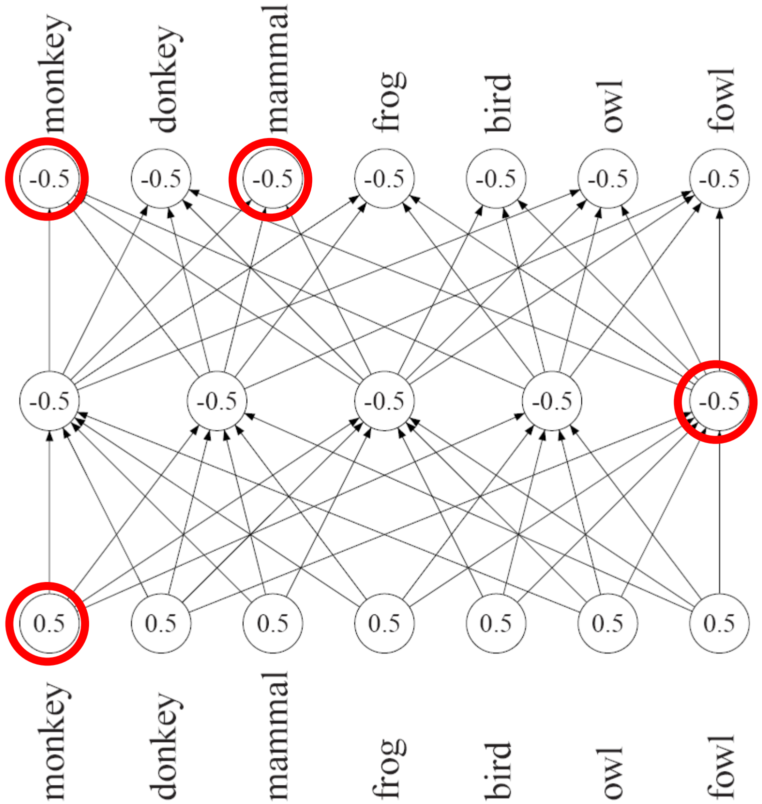
	monkey	donkey	mammal	frog	bird	owl	fowl
g_1					×		×
g_2					×	×	
g_3				×			
g_4		×	×				
g_5	×		×				

- set *G* constitutes middle layer
- blank cells of formal context constitute links in neural network

AIFB

Example

- as solved via FCA (antitone version):

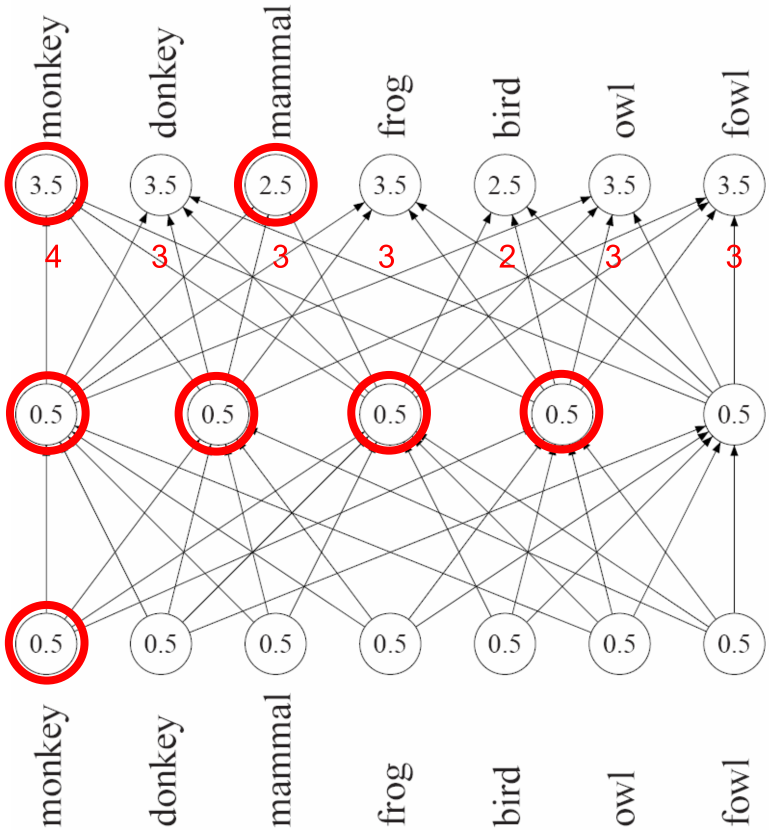


all links have weight -1

AIFB

Example

- as solved via FCA (monotone version):



all links have weight 1

AIFB

Conclusion

- neural networks defined via FCA calculate closure in one step (no recurrence needed)
- in some cases middle layer even smaller than using one middle layer neuron per implication

Directions for Future Work

- how to turn implication set into formal context without generating the whole lattice
  (because size might be exponential)
- extension to horn clauses
- how to use these results for the extraction part of the neural-symbolic learning cycle
  (not straightforward due to symmetry)

Thank you!