Exploring Implications and General Concept Inclusions with High Confidence

Daniel Borchmann

QuantLA Summer Workshop 2014

2014-10-06



http://www.sealifeconservation.org.au/adopt-an-animal/

Exploration by Confidence





Accept this GCI?



Accept this GCI? It depends



Accept this GCI?

It depends

True if all mammals are considered



Accept this GCI?

It depends

- True if all mammals are considered
- But exceptional, i. e. counterexamples rare



Accept this GCI?

It depends

- True if all mammals are considered
- But exceptional, i. e. counterexamples rare
- In general not desired



Accept this GCI?

It depends

- True if all mammals are considered
- But exceptional, i. e. counterexamples rare
- In general not desired

How to decide?



Accept this GCI?

It depends

- True if all mammals are considered
- But exceptional, i. e. counterexamples rare
- In general not desired

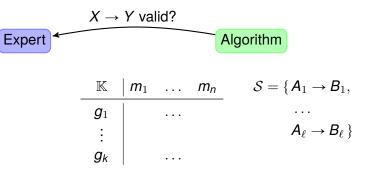
How to decide?

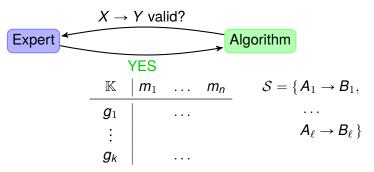
External source of information needed

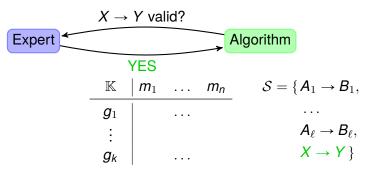


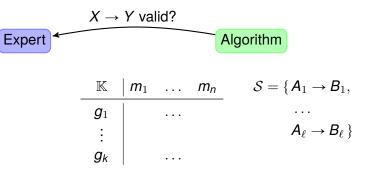
Algorithm

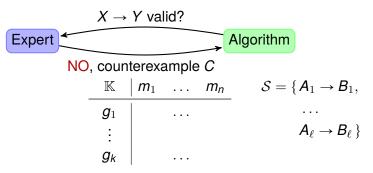
K	m_1	 m _n	$\mathcal{S} = \{ \mathbf{A}_1 \to \mathbf{B}_1, $
$oldsymbol{g}_1$			
:			$A_\ell o B_\ell$ }
g_k			

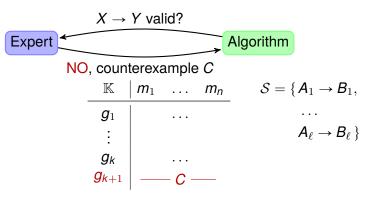








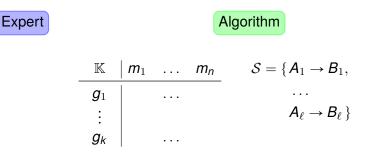




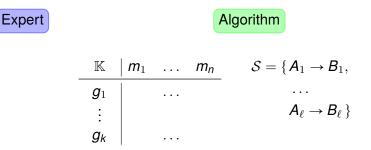


Algorithm

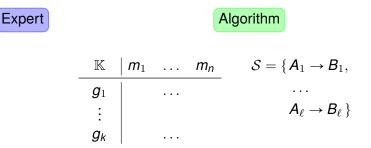
K	m_1	 m _n	$\mathcal{S} = \{ \mathbf{A}_1 \to \mathbf{B}_1, $
$oldsymbol{g}_1$			
:			$A_\ell o B_\ell$ }
g_k			



► Upon termination, Cn(S) = Th(K) and S is a base of the implicational knowledge of the expert



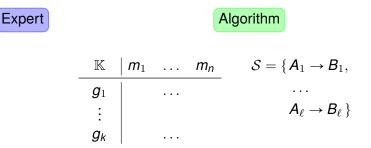
- ► Upon termination, Cn(S) = Th(K) and S is a base of the implicational knowledge of the expert
- Implications asked are always valid



- ► Upon termination, Cn(S) = Th(K) and S is a base of the implicational knowledge of the expert
- Implications asked are always valid

Problem

Data may contain errors



- ► Upon termination, Cn(S) = Th(K) and S is a base of the implicational knowledge of the expert
- Implications asked are always valid

Problem

Data may contain errors ~> ask implications with high confidence

Define

$$\operatorname{conf}_{\mathbb{K}}(X \to Y) := egin{cases} 1 & X' = arnothing \ rac{|(X \cup Y)'|}{|X'|} & ext{otherwise} \end{cases}$$

Define

$$\operatorname{conf}_{\mathbb{K}}(X \to Y) := egin{cases} 1 & X' = arnothing \ rac{|(X \cup Y)'|}{|X'|} & ext{otherwise} \end{cases}$$

For $\boldsymbol{c} \in [0,1]$ set

$$\mathsf{Th}_{c}(\mathbb{K}) := \{ X \to Y \mid \mathsf{conf}_{\mathbb{K}}(X \to Y) \ge c \}.$$

Define

$$\operatorname{conf}_{\mathbb{K}}(X \to Y) := egin{cases} 1 & X' = arnothing \ rac{|(X \cup Y)'|}{|X'|} & ext{otherwise} \end{cases}$$

For $\boldsymbol{c} \in [0, 1]$ set

$$\mathsf{Th}_{c}(\mathbb{K}) := \{ X \to Y \mid \mathsf{conf}_{\mathbb{K}}(X \to Y) \ge c \}.$$

Definition

Let *M* be a finite set. Then an *expert p* on *M* is a mapping

$$p\colon \operatorname{Imp}(M) \to \mathfrak{P}(M) \cup \{\top\}$$

such that

Define

$$\operatorname{conf}_{\mathbb{K}}(X \to Y) := egin{cases} 1 & X' = arnothing \ rac{|(X \cup Y)'|}{|X'|} & ext{otherwise} \end{cases}$$

For $\boldsymbol{c} \in [0,1]$ set

$$\mathsf{Th}_{c}(\mathbb{K}) := \{ X \to Y \mid \mathsf{conf}_{\mathbb{K}}(X \to Y) \ge c \}.$$

Definition

Let *M* be a finite set. Then an *expert p* on *M* is a mapping

$$p\colon \operatorname{Imp}(M) \to \mathfrak{P}(M) \cup \{\top\}$$

such that

•
$$p(X \rightarrow Y) = C \neq \top$$
 implies $X \subseteq C$ and $Y \nsubseteq C$,

Define

$$\mathsf{conf}_{\mathbb{K}}(X \to Y) := egin{cases} 1 & X' = arnothing \ rac{|(X \cup Y)'|}{|X'|} & ext{otherwise} \end{cases}$$

For $\boldsymbol{c} \in [0, 1]$ set

$$\mathsf{Th}_{c}(\mathbb{K}) := \{ X \to Y \mid \mathsf{conf}_{\mathbb{K}}(X \to Y) \ge c \}.$$

Definition

Let *M* be a finite set. Then an *expert p* on *M* is a mapping

$$p\colon \operatorname{Imp}(M) \to \mathfrak{P}(M) \cup \{\top\}$$

such that

▶
$$p(X \rightarrow Y) = C \neq \top$$
 implies $X \subseteq C$ and $Y \nsubseteq C$,

▶ $p(U \rightarrow V) = \top, p(X \rightarrow Y) = C \neq \top$ implies $U \oplus C$ or $V \subseteq C$.

To explore $\mathsf{Th}_c(\mathbb{K})$ with expert p with background knowledge $\mathcal S$ means to find a base of

 $\mathsf{Th}(\pmb{\rho}) \cap \mathsf{Th}_{\pmb{c}}(\mathbb{K})$

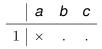
with background knowledge S,

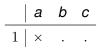
To explore $\mathsf{Th}_c(\mathbb{K})$ with expert p with background knowledge $\mathcal S$ means to find a base of

 $\mathsf{Th}(p) \cap \mathsf{Th}_{c}(\mathbb{K})$

with background knowledge $\mathcal{S},$ i. e. to compute a set $\mathcal{B} \subseteq \mathsf{Imp}(M)$ such that

$$Cn(\mathcal{B} \cup \mathcal{S}) = Cn(Th(\boldsymbol{p}) \cap Th_{\boldsymbol{c}}(\mathbb{K})).$$





Exploration by Confidence

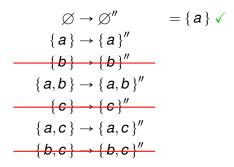
	а	b	С
1	×		

$$\begin{split} \varnothing &\to \varnothing'' \\ & \{a\} \to \{a\}'' \\ & \{b\} \to \{b\}'' \\ & \{a,b\} \to \{a,b\}'' \\ & \{c\} \to \{c\}'' \\ & \{a,c\} \to \{c,c\}'' \\ & \{b,c\} \to \{b,c\}'' \end{split}$$

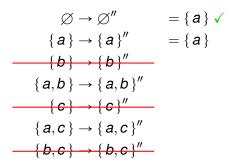
	а	b	С
1	×		

	а	b	С
1	×		

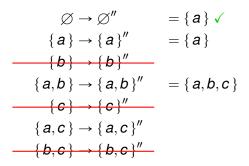
	а	b	С
1	×	•	



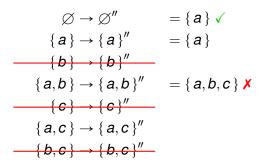
	а	b	С
$1 \mid$	×	•	



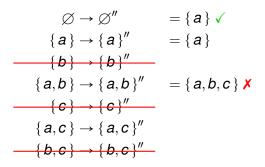
	а	b	С
1	×	•	



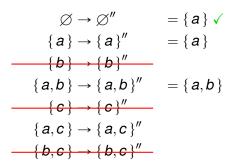
	а	b	С
1	×	•	



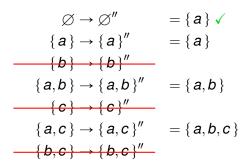
	а	b	С
1	×		
2	×	×	



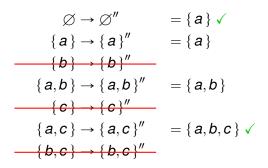
	а	b	С
1	×		
2	×	×	



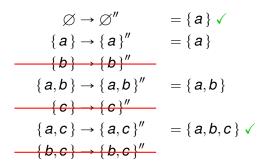
	а	b	С
1	×		
2	×	×	



	а	b	С
1	×		
2	×	×	



	а	b	С
1	× ×		
2	×	×	



Note

					()
		h	~	$\{a\} \rightarrow \{a\}''$	$= \{ a \}$
_	a			$\{b\} \rightarrow \{b\}''$	
	$\begin{array}{c c}1 & \times \\2 & \times\end{array}$	•	•	$\{a,b\} \rightarrow \{a,b\}''$	= { a , b }
	$2 \mid \times$	×	•	$ \{c\} \rightarrow \{c\}''$	
				$\{a,c\} \rightarrow \{a,c\}''$	= { <i>a</i> , <i>b</i> , <i>c</i>
				$\{b,c\}\rightarrow\{b,c\}''-$	

 $\emptyset \to \emptyset''$

 $= \{a\} \checkmark$

Note

► Order of premises extends ⊆

} 🗸

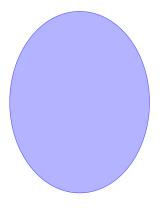
				$\varnothing \to \varnothing''$	={a}√
		Ь	~	$\{a\} \rightarrow \{a\}''$	$= \{ a \}$
	а			$\{b\} \rightarrow \{b\}''$	
1	× ×	•	•	$\{a,b\} \rightarrow \{a,b\}''$	$= \{ a, b \}$
2	×	Х	•	$\frac{\{c\} \rightarrow \{c\}''}{\{c\}}$	
				$\{a,c\} \rightarrow \{a,c\}''$	= { a , b , c } ✓
				$ \{b, c\} \rightarrow \{b, c\}''$	

. 11

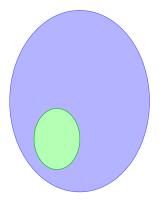
Note

- ► Order of premises extends ⊆
- Premises are always closed under known implications

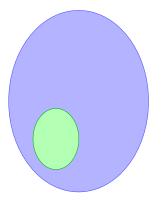
Exploration by Confidence



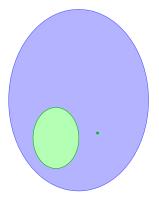
► Interesting Implications *L*



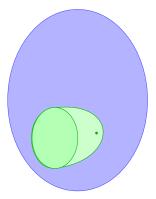
- ► Interesting Implications *L*
- ► Known Implications S



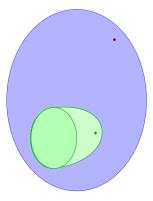
- ► Interesting Implications *L*
- Known Implications S



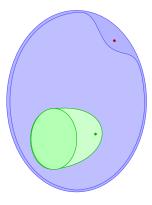
- ► Interesting Implications *L*
- Known Implications S
- Expert can *confirm* implications,



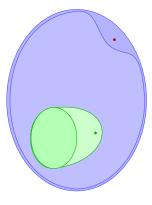
- ► Interesting Implications *L*
- Known Implications S
- Expert can *confirm* implications, extending S



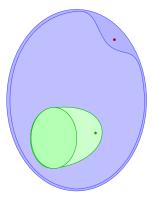
- ► Interesting Implications *L*
- Known Implications S
- Expert can *confirm* implications, extending S
- Expert can *reject* implications, providing counterexamples,



- ► Interesting Implications *L*
- Known Implications S
- Expert can *confirm* implications, extending S
- Expert can *reject* implications, providing counterexamples, shrinking L



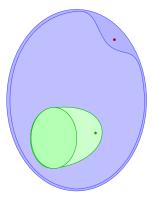
- ► Interesting Implications £
- Known Implications S
- Expert can *confirm* implications, extending S
- Expert can *reject* implications, providing counterexamples, shrinking L
- Iterate until $\mathcal{L} = \mathcal{S}$.



- ► Interesting Implications £
- Known Implications S
- Expert can *confirm* implications, extending S
- Expert can *reject* implications, providing counterexamples, shrinking L
- Iterate until $\mathcal{L} = \mathcal{S}$.

Which interesting but unknown implications are valid in our domain?

• Classical Attribute Exploration: $\mathcal{L} = \mathsf{Th}(\mathbb{K})$



- ► Interesting Implications £
- Known Implications S
- Expert can *confirm* implications, extending S
- Expert can *reject* implications, providing counterexamples, shrinking L
- Iterate until $\mathcal{L} = \mathcal{S}$.

- Classical Attribute Exploration: $\mathcal{L} = \mathsf{Th}(\mathbb{K})$
- Exploration by Confidence: $\mathcal{L} = \mathsf{Th}_{c}(\mathbb{K})$

Replace occurrence of $(\cdot)''$ with $\mathsf{Th}_{c}(\mathbb{K})(\cdot).$

```
Replace occurrence of (\cdot)'' with \mathsf{Th}_{c}(\mathbb{K})(\cdot).
```

Problems

```
Replace occurrence of (\cdot)'' with \mathsf{Th}_{c}(\mathbb{K})(\cdot).
```

Problems

It doesn't work,

```
Replace occurrence of (\cdot)'' with \mathsf{Th}_{c}(\mathbb{K})(\cdot).
```

Problems

It doesn't work, i. e. exploration is only "approximative":

Replace occurrence of $(\cdot)''$ with $\mathsf{Th}_c(\mathbb{K})(\cdot)$.

Problems

It doesn't work, i.e. exploration is only "approximative":

 $\mathsf{Th}(\rho) \cap \mathsf{Cn}(\mathsf{Th}_c(\mathbb{K})) \supseteq \mathsf{Cn}(\mathcal{B} \cup \mathcal{S}) \supsetneq \mathsf{Cn}(\mathsf{Th}(\rho) \cap \mathsf{Th}_c(\mathbb{K}))$

Replace occurrence of $(\cdot)''$ with $\mathsf{Th}_c(\mathbb{K})(\cdot)$.

Problems

It doesn't work, i. e. exploration is only "approximative":

 $\mathsf{Th}(\rho) \cap \mathsf{Cn}(\mathsf{Th}_{\boldsymbol{c}}(\mathbb{K})) \supseteq \mathsf{Cn}(\mathcal{B} \cup \mathcal{S}) \supsetneq \mathsf{Cn}(\mathsf{Th}(\rho) \cap \mathsf{Th}_{\boldsymbol{c}}(\mathbb{K}))$

• Closures under $Th_c(\mathbb{K})$ are (potentially) expensive

Idea

Instead of

$$X \to \mathsf{Th}_{c}(\mathbb{K})(X),$$

ask implications of the form

$$X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \}.$$

Idea

Instead of

$$X \to \mathsf{Th}_{c}(\mathbb{K})(X),$$

ask implications of the form

$$X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \}.$$

Problem

Doesn't work either.

•
$$c := \frac{1}{2}$$

- $S := \{ \{ a \} \rightarrow \{ b \} \}$
- ▶ *p* constantly ⊤



•
$$c := \frac{1}{2}$$

- $S := \{ \{ a \} \rightarrow \{ b \} \}$
- ▶ *p* constantly ⊤

Then

 $\text{conf}_{\mathbb{K}}(\{\,a\,\} \to \{\,c\,\}) \geqslant c,$

	a	b	С
1	×	×	
2	×	Х	
3	×	Х	
4	×	Х	Х
5	×	Х	×
6	×		×
7	×		Х
8			
9			
10			

_

•
$$c := \frac{1}{2}$$

- $S := \{ \{ a \} \rightarrow \{ b \} \}$
- ▶ *p* constantly ⊤

Then

$$\text{conf}_{\mathbb{K}}(\{\,a\,\} \to \{\,c\,\}) \geqslant \textit{c},$$

but $\{a\}$ is not closed,

	a	b	С
1	×	×	
2	×	Х	
3	×	Х	
4	×	Х	×
5	×	Х	\times
6	×		×
7	×		\times
8			
9			
10			

-

•
$$c := \frac{1}{2}$$

- $S := \{ \{ a \} \rightarrow \{ b \} \}$
- ▶ *p* constantly ⊤

Then

$$\text{conf}_{\mathbb{K}}(\{\,a\,\} \to \{\,c\,\}) \geqslant c,$$

but $\{\,a\,\}$ is not closed, and

$$\operatorname{conf}_{\mathbb{K}}(\{\mathbf{b}\} \to \{\mathbf{c}\}) = \frac{2}{5} < \mathbf{c},$$
$$\operatorname{conf}_{\mathbb{K}}(\{\mathbf{a},\mathbf{b}\} \to \{\mathbf{c}\}) = \frac{2}{5} < \mathbf{c},$$
$$\operatorname{conf}_{\mathbb{K}}(\emptyset \to \{\mathbf{c}\}) = \frac{4}{10} < \mathbf{c}.$$

	a	b	С
1	×	×	
2	×	Х	
3	×	Х	
4	×	Х	×
5	×	Х	×
6	×		×
7	×		×
8			
9			
10	.		

Ensure: when

$$X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \}$$

is asked in iteration *i*,

Ensure: when

$$X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \}$$

is asked in iteration *i*, then for $Y'' \subsetneq X$

$$\operatorname{conf}_{\mathbb{K}}(Y'' \to \{n\}) \ge c \implies \mathcal{B}_i \cup \mathcal{S}_i \models (Y'' \to \{n\}).$$

Ensure: when

$$X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \}$$

is asked in iteration *i*, then for $Y'' \subsetneq X$

$$\operatorname{conf}_{\mathbb{K}}(Y'' \to \{n\}) \ge c \implies \mathcal{B}_i \cup \mathcal{S}_i \models (Y'' \to \{n\}).$$

Types of Implications asked

Ensure: when

$$X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \}$$

is asked in iteration *i*, then for $Y'' \subsetneq X$

$$\operatorname{conf}_{\mathbb{K}}(Y'' \to \{n\}) \ge c \implies \mathcal{B}_i \cup \mathcal{S}_i \models (Y'' \to \{n\}).$$

Types of Implications asked

• $X \to X''$, where X is closed under $\mathcal{B}_i \cup \mathcal{S}_i$, but not an intent

Solution

Ensure: when

$$X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \}$$

is asked in iteration *i*, then for $Y'' \subsetneq X$

$$\mathsf{conf}_{\mathbb{K}}(Y'' \to \{n\}) \geqslant c \implies \mathcal{B}_i \cup \mathcal{S}_i \models (Y'' \to \{n\}).$$

Types of Implications asked

- $X \to X''$, where X is closed under $\mathcal{B}_i \cup \mathcal{S}_i$, but not an intent
- ▶ $X \to \{ m \in M \mid \operatorname{conf}_{\mathbb{K}}(X \to \{ m \}) \ge c \} \setminus (\mathcal{B}_i \cup \mathcal{S}_i)(X)$, where X is an intent

	а	b	с
1	×	×	
2	Х	×	
3	×	Х	
4	×	Х	×
5	×	Х	×
6	Х		×
7	Х		×
8			•
9			
10	•	•	•

•
$$c := \frac{1}{2}$$

• $S = \{\{a\} \rightarrow \{b\}\}\$
• p constantly \top

	а	b	с
1	×	×	
2	Х	Х	
3	×	Х	
4	Х	Х	×
5	Х	Х	Х
6	Х		Х
7	Х		Х
8	•		
9	•		
10	•		

• $c := \frac{1}{2}$
$\blacktriangleright S = \{\{a\} \rightarrow \{b\}\}$
• <i>p</i> constantly \top
Ø
{ a }
{ b }
{ a, b }
{ c }
$\{a, c\}$
$\set{b,c}$

	а	b	с
1	×	×	
2	Х	Х	•
3	×	Х	
4	×	Х	×
5	Х	Х	×
6	Х		Х
7	Х		Х
8	•		•
9			
10	•		•

• $c := \frac{1}{2}$
$\blacktriangleright S = \{\{a\} \rightarrow \{b\}\}$
• p constantly $ op$
Ø
{ a }
{ b }
{ a , b }
{ c }
{ a, c }
$\set{b,c}$

	а	b	С	
1	×	×		
2	Х	Х		
3	Х	Х		
4	×	Х	×	
5	×	Х	×	
6	Х		×	
$\overline{7}$	Х		×	
8			•	
9				
10	•	•	•	

Þ	${m c}:=rac{1}{2}$
Þ	$\mathcal{S} = \{ \{ a \} \rightarrow \{ b \} \}$
Þ	p constantly $ op$
	$\varnothing \rightarrow \{a\}$
	{ a }
	{ b }
	{ a, b }
	{ c }
	{ a, c }
	{ b, c }

	а	b	С
1	×	×	
2	Х	×	
3	×	Х	
4	Х	Х	×
5	Х	Х	×
6	Х	•	×
7	Х	•	×
8	•	•	•
9	•	•	•
10	•	•	•

	а	b	с
1	×	×	
2	×	Х	
3	×	Х	
4	×	×	×
5	×	Х	×
6	×		×
7	×		×
8	•		
9	•		
10	•		

	а	b	с
1	×	×	
2	×	Х	
3	×	Х	
4	×	×	×
5	×	Х	\times
6	×		\times
7	×		\times
8	•		
9	•		
10	•		•

$$c := \frac{1}{2}$$

$$S = \{\{a\} \rightarrow \{b\}\}$$

$$p \text{ constantly } \top$$

$$\bigotimes \rightarrow \{a\} \checkmark$$

$$\{a\} \rightarrow \{a, b, c\}$$

$$\{a, b\}$$

$$\{a, b\}$$

$$\{c\}$$

$$\{a, c\}$$

$$\{b, c\}$$

•

	а	b	С
1	×	×	
2	Х	×	
3	×	Х	
4	Х	Х	×
5	Х	Х	×
6	Х		×
7	Х		×
8			
9			
10	•	•	

►

	а	b	С
1	×	×	
2	Х	×	
3	×	Х	
4	×	Х	×
5	×	Х	×
6	Х		×
7	Х		×
8			
9			
10	•	•	

• $C := \frac{1}{2}$ $\flat \mathcal{S} = \{\{a\} \rightarrow \{b\}\}\}$ • p constantly \top $\emptyset \rightarrow \{a\} \checkmark$ $\{a\} \rightarrow \{a, b, c\} \checkmark$ {b} $\{a, b\}$ {c} $\{a, c\}$ $\{b, c\}$

Used attribute exploration to handle exceptional counterexamples

- Used attribute exploration to handle exceptional counterexamples
- Generalized attribute exploration

- Used attribute exploration to handle exceptional counterexamples
- Generalized attribute exploration
- Developed "Exploration by Confidence"

- Used attribute exploration to handle exceptional counterexamples
- Generalized attribute exploration
- Developed "Exploration by Confidence"

Outlook to GCIs

- Used attribute exploration to handle exceptional counterexamples
- Generalized attribute exploration
- Developed "Exploration by Confidence"

Outlook to GCIs

Same ideas can be applied

- Used attribute exploration to handle exceptional counterexamples
- Generalized attribute exploration
- Developed "Exploration by Confidence"

Outlook to GCIs

- Same ideas can be applied
- Counterexamples need to be connected subinterpretations

- Used attribute exploration to handle exceptional counterexamples
- Generalized attribute exploration
- Developed "Exploration by Confidence"

Outlook to GCIs

- Same ideas can be applied
- Counterexamples need to be connected subinterpretations
- Extra expert interaction required (due to growing set of "attributes")

Fact?

$\textbf{Venomous} \sqcap \textbf{Bird} \sqsubseteq \bot$