Learning Datalog Programs from Input and Output

Yisong Wang, Xu Wang, and Yi Huang
Department of Computer Science, Guizhou University
Can we learn from this incomplete scenario?

With knowing none of rule or a few (not the whole) rules.

......
Inductive learning task

• Given: a background theory (Datalog program) $B$ and a set $E$ of examples, which are in the form of a pair $\langle I, O \rangle$ where input is $I$ and output is $O$.

• Find: a hypothesis $H$ such that for every example $\langle I, O \rangle$ in $E$, $O$ is the least (Herbrand) model of $B \cup H \cup I$.

Note:\ The learning task is denoted by $ILT(B, E)$. A hypothesis $H$ of the learning task $ILT(B, E)$ is called a solution to $ILT(B, E)$.
Condition for existence of solution

Given an inductive learning task $\text{ILT}(B, E)$, there exists a solution $H$ to $\text{ILT}(B, E)$ if and only if $E$ is coherent and $E$ is consistent w.r.t. $B$.

• A set $E$ of examples is coherent if
  • $I \subseteq O$ for every $\langle I, O \rangle \in E$, and
  • each two distinct examples $\langle I_1, O_1 \rangle$ and $\langle I_2, O_2 \rangle$ in $E$ satisfy monotonicity and convergence.
    • monotonicity: $I_1 \subseteq I_2$ implies $O_1 \subseteq O_2$
    • convergence: $I_1 \subseteq O_2$ implies $O_1 \subseteq O_2$

• A set $E$ of examples is consistent w.r.t. a background theory $B$ if $O \models B$ for every $\langle I, O \rangle$ in $E$. 
Algorithm 1: \(LFIO(B, E)\)

Input: \(E\): a set of observations, \(B\): a background theory

Output: A Datalog program \(P\)

1. \(P \leftarrow \emptyset\);
2. \textbf{foreach} \(\langle I, O \rangle \in E\) \textbf{do}
3. \hspace{1em} \(M \leftarrow\) the least model of \(B \cup P\);
4. \hspace{1em} \textbf{foreach} \(p \in O \setminus (M \cup I)\) \textbf{do}
5. \hspace{2em} \text{Let} \(r\) be the rule \(p \leftarrow (I \setminus M)\);
6. \hspace{2em} \textbf{if} \(\neg \exists r' \in B \cup P\) such that \(r' \preceq r\) \textbf{then}
7. \hspace{3em} \text{Remove each rule} \(r''\) with \(r \preceq r''\) from \(P\);
8. \hspace{3em} \(P \leftarrow P \cup \{r\}\);
9. \hspace{2em} \textbf{end}
10. \hspace{1em} \textbf{end}
11. \textbf{end}
12. \textbf{return} \(P\);
A rule $r$ subsumes a rule $r'$ if there exists a substitution $\theta$ such that $hd(r)\theta = hd(r')$ and $bd(r)\theta \subseteq bd(r')$. We denote $hd(r)$ is the head of rule $r$ and $bd(r)$ is the set of bodies of rule $r$. A substitution $\theta$ is a function from variables to terms and The application $e\theta$ of $\theta$ to an expression $e$ replace all occurrences of each variable in $e$ with the same term.
Soundness of the algorithm \textit{LFIO}

If $E$ is \textbf{coherent} and it is \textbf{consistent} w.r.t. $B$ then $O$ is the least model of $B \cup LFIO(B, E) \cup I$ for every $\langle I, O \rangle \in E$.

\textbf{Note}: We assume that the given examples in $E$ are coherent and are consistent w.r.t. the background theory $B$ in the inductive learning task in what follows.
Properties of the Algorithm \textit{LFIO}

• **Modular:** Given $E = E_1 \cup E_2$ be a set of examples and a background theory $B$. Let $P_1 = LFIO(B, E_1)$ and $P_2 = LFIO(B, E_2)$. Then $O$ is the least model of $B \cup P_1 \cup P_2 \cup I$ for every $\langle I, O \rangle \in E$.

• **Incremental:** Given $E = E_1 \cup E_2$ be a set of examples and a background theory $B$ and $E_1 \cap E_2 = \emptyset$. Let $P_1 = LFIO(B, E_1)$ and $P_2 = LFIO(B \cup P_1, E_2)$. Then $O$ is the least model of $B \cup P_2 \cup I$ for every $\langle I, O \rangle \in E$. 
Example

Let $B = \emptyset$ and $E$ consists of $\langle \emptyset, \emptyset \rangle$, $\langle \{p\}, \{p\} \rangle$, $\langle \{q\}, \{q\} \rangle$, $\langle \{r\}, \{p, r\} \rangle$, $\langle \{p, q\}, \{p, q, r\} \rangle$, $\langle \{q, r\}, \{p, q, r\} \rangle$, $\langle \{p, r\}, \{p, r\} \rangle$, $\langle \{p, q, r\}, \{p, q, r\} \rangle$.

<table>
<thead>
<tr>
<th>step</th>
<th>$\langle I, O \rangle$</th>
<th>$r$</th>
<th>ID</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\langle \emptyset, \emptyset \rangle$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$\langle {p}, {p} \rangle$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$\langle {q}, {q} \rangle$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$\langle {r}, {p, r} \rangle$</td>
<td>$p \leftarrow r$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>$\langle {p, q}, {p, q, r} \rangle$</td>
<td>$r \leftarrow p, q$</td>
<td>2</td>
<td>1,2</td>
</tr>
<tr>
<td>6</td>
<td>$\langle {q, r}, {p, q, r} \rangle$</td>
<td>$p \leftarrow q, r$</td>
<td>3</td>
<td>1,2</td>
</tr>
<tr>
<td>7</td>
<td>$\langle {p, r}, {p, r} \rangle$</td>
<td></td>
<td></td>
<td>1,2</td>
</tr>
<tr>
<td>8</td>
<td>$\langle {p, q, r}, {p, q, r} \rangle$</td>
<td></td>
<td></td>
<td>1,2</td>
</tr>
</tbody>
</table>
Properties of the Algorithm **LFIO**

- **Simplified**
  - In the case $I = O$ for some $\langle I, O \rangle \in E$, no rule is added in the term of LFIO and
  - In the case two different examples $\langle I_1, O_1 \rangle$ and $\langle I_2, O_2 \rangle \in E$ and $O_1 \setminus I_1 = O_2 \setminus I_2$, if $I_1 \subseteq I_2$, then for each rule which is produced in line 5 of the algorithm by $\langle I_2, O_2 \rangle$, it must be subsumed by one rule which is produced by $\langle I_1, O_1 \rangle$. 
Example

Let $B = \emptyset$ and $E$ consists of $\langle \emptyset, \emptyset \rangle$, $\langle \{p\}, \{p\} \rangle$, $\langle \{q\}, \{q\} \rangle$, $\langle \{r\}, \{p, r\} \rangle$, $\langle \{p, q\}, \{p, q, r\} \rangle$, $\langle \{q, r\}, \{p, q, r\} \rangle$, $\langle \{p, r\}, \{p, r\} \rangle$, $\langle \{p, q, r\}, \{p, q, r\} \rangle$.

We have that $E$ can be simplified like the set $\langle \{r\}, \{p, r\} \rangle$, $\langle \{p, q\}, \{p, q, r\} \rangle$.

<table>
<thead>
<tr>
<th>step</th>
<th>$\langle I, O \rangle$</th>
<th>$r$</th>
<th>ID</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\langle {r}, {p, r} \rangle$</td>
<td>$p \leftarrow r$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$\langle {p, q}, {p, q, r} \rangle$</td>
<td>$r \leftarrow p, q$</td>
<td>2</td>
<td>1,2</td>
</tr>
</tbody>
</table>
Summary

• The output is the least (Herbrand) model of a solution to an inductive learning task together with its background theory and input.

• A modular and incremental inductive learning algorithm was presented for this inductive learning task.
Acknowledgement

We thank reviewers for their helpful comments. This work is partially supported by NSFC under grant 63170161, Stadholder Fund of Guizhou Province under grant (2012)62, Outstanding Young Talent Training Fund of Guizhou Province under grant (2015)01 and Science and Technology Fund of Guizhou Province under grant [2014]7640.
References


