Integrating Logical Reasoning with Probabilistic Chain Graphs

Arjen Hommersom, Nivea Ferreira, and Peter Lucas
{arjenh,nivea,peterl}@cs.ru.nl

Institute for Computing and Information Sciences
Radboud University Nijmegen
The Netherlands
Motivation

- First-order logic: good for relational reasoning in various ways about classes of objects
- Probabilistic graphical models: good for reasoning with uncertainty

⇒ why not combining them?
- Markov logic (generates Markov networks)
- Bayesian logic programs (generates Bayesian networks)
- Probabilistic Horn logic (abductive Bayesian-network reasoning)
- Chain logic (abductive chain-graph reasoning)
Chain Graphs

- Graphical representation associated with a Bayesian network is **not unique**
  - different graphs may represent the same independence information

- Markov networks can be seen as the weakest type of graphical models
  - much of the subtleties of representing conditional dependence and independence cannot be handled

- Unique **chain graph** representatives of Bayesian networks (**essential graphs**)
  - Bayesian networks and Markov networks as special cases
Chain Graph Definition

- A chain graph is a hybrid graph with the restriction that no directed cycles exist.

- Factorisation: chain graphs can be interpreted as an acyclic directed graph of chain components.

\[
P(X_V) = \prod_{C \in C} P(X_C \mid X_{pa(C)})
\]

with \( V = \bigcup_{C \in C} C \), and where each \( P(X_C \mid X_{pa(C)}) \) factorises according to

\[
P(X_C \mid X_{pa(C)}) = Z^{-1}(X_{pa(C)}) \prod_{M \in M(C)} \varphi_M(X_M)
\]
Chain Graph Example

Influenza ($I$) causes coughing ($C$), where coughing is known as a possible cause for hoarseness ($H$). In addition, coughing is known to be associated with dyspnoea (shortness of breath) ($D$). Dyspnoea restricts the oxygen supply to the blood circulation; the resulting low oxygen saturation of the blood will turn the skin to colour blue ($B$)
Horn Clauses

- A formula in first-order logic
- A Horn-clause has a general form given by

\[ A \leftarrow B_1, \ldots, B_n \]

where \( A \) is the head and \( B_1, \ldots, B_n \) the body of the clause.

- Reasoning:
  - standard model-theoretic semantics, defined in terms of the logical consequence operator \( \models \)
  - procedural semantics, defined in terms of the deduction relation \( \vdash \)
Abduction Logic

Horn clauses of the form:

\[ D \leftarrow B_1, \ldots, B_n : R_1, \ldots, R_m \]

where

- \( D \): head of the clause, a predicate or \( \perp \)
- \( B_1, \ldots, B_n \): body of the clause, a set of predicates (*will become ‘random variables’*)
- \( R_i \): templates, to express relations between variables

Both the ‘,’ as well as the ‘:’ are interpreted as a conjunction.
Influenza: Logical Specification

\[
\begin{align*}
I(x) & \leftarrow: \varphi_I(x) \\
C(x) & \leftarrow I(y): \varphi_{C,I}(x,y), \varphi_{C,D}(x,z) \\
D(x) & \leftarrow I(y): \varphi_{C,I}(z,y), \varphi_{C,D}(z,x) \\
H(x) & \leftarrow C(y): \varphi_{H,C}(x,y) \\
B(x) & \leftarrow D(y): \varphi_{B,D}(x,y) \\
\bot & \leftarrow \varphi_{C,I}(x,y), \varphi_{C,D}(\bar{x},z)
\end{align*}
\]

where the \( \varphi \)s are relations \( R_k \)
Reasoning: Explanations

Let:
- \( T \): an abductive theory, which is a set of formulae
- \( A \): the set of all assumables
- \( A' \): denote the set of ground instances of \( A \)

An explanation \( E \) of a set of observations \( O \) based on the pair \( \langle T, A \rangle \) is defined as a set of ground assumables \( E \subseteq A' \) satisfying the following conditions:
- \( T \cup E \models O \), and
- \( T \cup E \) is consistent, i.e., \( T \cup E \nvdash \bot \).
Chain Logic Syntax

Syntax of chain logic consists of:
- Formulae in abduction logic
- Weight declarations, which are of the form
  \[ weight(a_1 : w_1, \ldots, a_n : w_n) \]
  where \( a_i \) represents an atom and \( w_i \) real, such that a weight declaration contains all instances of a predicate

Then, we define:
- Assumables \( A \): atoms that occur in \( weight \)
- Hypothesis \( H \): consistent set of ground atoms in \( weight \) (one per \( weight \))
**Influenza**

Potential functions:

<table>
<thead>
<tr>
<th>( \varphi_{CI} )</th>
<th>( i )</th>
<th>( \bar{i} )</th>
<th>( \varphi_{CD} )</th>
<th>( d )</th>
<th>( \bar{d} )</th>
<th>( \varphi_{HC} )</th>
<th>( c )</th>
<th>( \bar{c} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c )</td>
<td>8</td>
<td>2</td>
<td>( c )</td>
<td>18</td>
<td>2</td>
<td>( h )</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>( \bar{c} )</td>
<td>1</td>
<td>10</td>
<td>( \bar{c} )</td>
<td>5</td>
<td>2</td>
<td>( \bar{h} )</td>
<td>0.4</td>
<td>0.9</td>
</tr>
</tbody>
</table>

The abduction clauses:

\[ I(x) \leftarrow: \varphi_I(x) \]
\[ C(x) \leftarrow I(y) : \varphi_{C,I}(x,y), \varphi_{C,D}(x,z) \]
\[ D(x) \leftarrow I(y) : \varphi_{C,I}(z,y), \varphi_{C,D}(z,x) \]
\[ H(x) \leftarrow C(y) : \varphi_{H,C}(x,y) \]
\[ B(x) \leftarrow D(y) : \varphi_{B,D}(x,y) \]
\[ \bot \leftarrow \varphi_{C,I}(x,y), \varphi_{C,D}(\bar{x},z) \]

Weights of the assumables \( weight(\varphi_{CD}(t,t) : 18, \varphi_{CD}(t,f) : 2, \varphi_{CD}(f,t) : 5, \varphi_{CD}(f,f) : 2) \)
Chain Logic Semantics

Abductive theory:
\[ T = \{ I(x) \leftarrow: \varphi_I(x), \]
\[ C(x) \leftarrow I(y): \varphi_{C,I}(x,y), \varphi_{C,D}(x,z), \]
\[ D(x) \leftarrow I(y): \varphi_{C,I}(z,y), \varphi_{C,D}(z,x), \]
\[ H(x) \leftarrow C(y): \varphi_{H,C}(x,y), \]
\[ B(x) \leftarrow D(y): \varphi_{B,D}(x,y), \]
\[ \bot \leftarrow \varphi_{C,I}(x,y), \varphi_{C,D}(\bar{x},z) \} \]

where each of the variables has \( \{f, t\} \) as domain

It now holds that:
\[ T \cup E \models H(t) \text{ and } T \cup E \not\models \bot, \text{ with} \]
\[ E = \{ \varphi_I(t), \varphi_{H,C}(t,t), \varphi_{C,I}(t,t), \varphi_{C,D}(t,t) \} \]
Minimal Explanations

A **minimal explanation** $E$ of $O$ is an explanation whose proper subsets are not explanations of $O$. The set of all minimal explanations is denoted by $\mathcal{E}_T(O)$

Suppose we would like to calculate if a person is blue, i.e., $P(B(t))$; we obtain the minimal explanations for $B(t)$, i.e., $\mathcal{E}_T(B(t))$, as the set with the following 8 members:

$$\begin{align*}
\{ \varphi_{B,D}(t,t), \varphi_{C,D}(t,t), \varphi_{C,I}(t,t), \varphi_I(t) \} \\
\{ \varphi_{B,D}(t,t), \varphi_{C,D}(t,t), \varphi_{C,I}(t,f), \varphi_I(f) \} \\
\vdots
\end{align*}$$

$$P(B(t)) = \sum_{E \in \mathcal{E}_T(B(t))} P(E) = 27.7/Z \approx 0.24$$
Probabilities of Formulae

Suppose $E$ is a minimal explanation. Then, given $T$, $P_T(E)$ is obtained by marginalisation:

$$P_T(E) = P_T(\bigvee H_i) = \sum_i P_T(H_i)$$

as $H_i$’s are mutually exclusive hypotheses (one atom per weight)

**Theorem** If $\mathcal{E}_T(\psi)$ is the set of minimal explanations of the conjunction of atoms $\psi$ from the chain logic theory $T$, then:

$$P_T(\psi) = \sum_{E \in \mathcal{E}_T(\psi)} P_T(E)$$
Reasoning: Abductively

Direction of reasoning about $B$: from $B$ to $I$, but ignoring $H$

Probabilistic reasoning = logical reasoning
Final Considerations

- Chain logic is inspired by Poole’s probabilistic Horn logic

- We present here a language that can be used for the specification of both Bayesian and Markov network models

- Maintaining a close relation between logical and probabilistic reasoning – without loss of expressiveness

- Learning of parameters by exploiting the relation between chain graphs and chain logic