Counterterrorism
Reasoning About Rare Events

Tod S. Levitt
Information Extraction & Transport, Inc.
Arlington, Virginia

September 20, 2007
Rare Events: Philosophical Issues

• Is “p(terrorist-event | evidence)” an admissible expression?

• If there can be such a distribution, how might it be estimated?

• If it is estimable, what experiment can be done to sample it?
Rare Events: Philosophical Answers

- **Is “p(terrorist-event | evidence)” an admissible expression?**
  - Classical: no; Bayesian: yes
  - See L.J. Savage; Foundations of Statistics; 1954

- **If there can be such a distribution, how might it be estimated?**
  - Evidential accrual via hierarchical Bayesian inference
  - Posterior probability at level n is the prior probability for level n+1

- **If it is estimable, what experiment can be done to sample it?**
  - Suitably constrain the event space over actors, targets and durations
  - Note Ramsey–De Finetti theorem: Bayesian updating converges to distribution when sampling from an exchangeable event space
Analyst Reasoning Issues

• Analysts do not usually attempt to estimate $p(\text{terrorist-event} \mid \text{evidence})$

• Leads to behaviors as if underestimating $p(\text{terrorist-event}:=\text{false} \mid \text{evidence})$

• Given terrorist attack hypothesis there is a tendency to search only for confirming evidence

• This leads to over-focus on specific hypotheses and missed hypothesis opportunities

• Our objective is a mathematical basis and computable approach to estimate $p(\text{terrorist-event} \mid \text{evidence})$
The Madrid Railway Attack, March 2004

- Madrid Cell Leader: 1
- Bomb Carriers: 14
- Explosives Acquisition: 6
- Bomb Maker: 1
- Communication between cell leader and members: face-to-face, mobile phones, and internet.

There are indications the AQ-Europe courier conveyed data about the operation to high-ranking AQ members.
CT Analysis: Link Diagrams

actors

situated
acts

actor

event
Links Sample Space Concept

![Diagram showing the concept of links in a sample space with different levels of association.](image)

- terrorist actor
- cell members
- supporters
- sympathizers
- social acquaintances
- public contacts

- vendors
- landlords
- officials
- workplace acquaintances
- university students
- typical mosque members
- Muslim-NGO members
- Salafi mosque members
- family members

exponentially increasing probability of association
From Links to Chains

Linked:

Inducted:

Causal:
Link-Direct Evidential Accrual

Hypothesis:

Evidence:

Objective: Quantify belief in attack given evidence

Methodology: Bayesian probabilistic evidential accrual
Enabling Reasoning Belief Quantification

**Hypothesis H**: Terrorists executing attack plan on site
- States: true, false

**Evidence E**: Surveillance at site
- States: true, false

**Bayes Rule**

\[ p(H|E) = \frac{p(E|H)p(H)}{p(E)} \]

\[ p(\ E=\text{true} \mid H=\text{true} \) = .999 \]
\[ p(\ E=\text{true} \mid H=\text{false} \) = .001 \]
\[ p(\ E=\text{false} \mid H=\text{true} \) = .001 \]
\[ p(\ E=\text{false} \mid H=\text{false} \) = .999 \]

\[ L_E = \frac{p(E|H=\text{true})}{p(E|H=\text{false})} \]

\[ L_E=\text{true} = .999/.001 = 999 \]
\[ L_E=\text{false} = .001/.999 = .001001 \]
Rare Versus Weak Evidence

- **Weak Evidence**: $L_E = \frac{p(E|H=true)}{p(E|H=false)}$ is close to 1
- **Rare Evidence**: $p(E|H=true)$ and $p(E|H=false)$ are close to 0

- Weak evidence is not necessarily rare
- Rare evidence is not necessarily weak

- Weak evidence inference order doesn’t matter
- Rare evidence inference order matters a lot
# Rare & Weak Evidence Examples

Consider $p(\text{Presence Pattern} \mid \text{Terrorist Surveillance})$

<table>
<thead>
<tr>
<th>Presence Pattern</th>
<th>Explanation</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>twice in one week</td>
<td>tourist</td>
<td>not_rare &amp; weak</td>
</tr>
<tr>
<td>once a month for a year</td>
<td>monthly lecture</td>
<td>rare &amp; weak</td>
</tr>
<tr>
<td>seven days in a row</td>
<td>terrorist</td>
<td>rare &amp; not_weak</td>
</tr>
<tr>
<td>five days in a row</td>
<td>security guard</td>
<td>not_rare &amp; not_weak</td>
</tr>
</tbody>
</table>
Hypothesis:

\[ p(\text{attack} | t & m & r & i & s) = \frac{p(t & m & r & i & s | \text{attack}) p(\text{attack})}{p(t & m & r & i & s)} = \frac{p(t|\text{attack}) p(m|\text{attack}) p(r|\text{attack}) p(i|\text{attack}) p(s|\text{attack})}{p(t & m & r & i & s)} \]
Link-Direct Evidential Accrual Issues

Hypothesis:

Evidence:

Prior Conditionals: \( p(\text{evidence} \mid \text{attack}) \) not reasonably estimable

Other Issues:

- \( p(\text{telephone\_conversation} \mid \text{attack}:\text{false}) \) is high, but analyst likely ranks low
- Exchangeability of events assures that order of evidence arrival does not affect result of Bayesian updating in the limit of infinite pieces of evidence (Ramsey-DeFinetti theorem)
- But ordering update effects are important in finite sequences of evidence
# Rare Evidence Order Sensitive

<table>
<thead>
<tr>
<th></th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(i</td>
<td>A)$</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td><strong>0.099</strong></td>
</tr>
<tr>
<td>$p(i</td>
<td>A^c)$</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td>$p(r</td>
<td>A)$</td>
<td>0.999</td>
<td>0.999</td>
<td><strong>0.099</strong></td>
<td>0.999</td>
</tr>
<tr>
<td>$p(r</td>
<td>A^c)$</td>
<td>0.001</td>
<td>0.001</td>
<td><strong>0.001</strong></td>
<td>0.001</td>
</tr>
<tr>
<td>$p(m</td>
<td>A)$</td>
<td>0.999</td>
<td>0.999</td>
<td><strong>0.099</strong></td>
<td>0.999</td>
</tr>
<tr>
<td>$p(m</td>
<td>A^c)$</td>
<td>0.001</td>
<td>0.001</td>
<td><strong>0.001</strong></td>
<td>0.001</td>
</tr>
<tr>
<td>$p(t</td>
<td>A)$</td>
<td><strong>0.099</strong></td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>$p(t</td>
<td>A^c)$</td>
<td><strong>0.001</strong></td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$L_s (A)$</td>
<td><strong>25.12</strong></td>
<td><strong>6.38</strong></td>
<td><strong>5.38</strong></td>
<td><strong>4.09</strong></td>
<td><strong>25.12</strong></td>
</tr>
</tbody>
</table>
Structuring Evidence for Accrual

Evidence O: Repeat observation of suspect person  
- States: true, false

Evidence R: Reliability of report  
- States: high, neutral, suspect

Evidence E: Surveillance at site  
- States: true, false

Evidence D: Site is a terrorist target  
- States: true, false

Hypothesis H: Terrorists executing attack plan on site  
- States: true, false
Hierarchical Dependencies

Evidence O: Repeat observation of suspect person
- States: true, false

Evidence R: Reliability of report
- States: highly, neutral, suspect

Evidence E: Surveillance at site
- States: true, false

Evidence D: Site is a terrorist target
- States: true, false

Hypothesis H: Terrorists executing attack plan on site
- States: true, false
Hierarchical Bayesian Inference Issues

- Must account for multiple types of uncertainty
- Conditional independence cannot necessarily be assumed – variable relations form a directed acyclic graph
- Evidence arrives asynchronously in arbitrary order
CT Analysis: Causal Reasoning
Causal Reasoning Issues

• Scientific inference strongly discouraged causal reasoning throughout the 20th century

• Breakthroughs in Bayesian inference support the principled use of causal scientific reasoning

• Evidential accrual algorithms provide mathematical foundation for well-structured inductive causal inference
Bayesian Networks Nano-Tutorial

Model 1:
\[ p(X, Y, Z) = p(Z | X,Y) p(Y | X) p(X) \]

Model 2:
\[ p(X, Y, Z) = p(Z | X,Y) p(Y) p(X) \]

Universal Solutions:
\[ p(Z) = \sum_{X,Y} p(X,Y,Z) = \sum_{X,Y} p(Z | X,Y) p(Y) p(X) \]

Compare Models:
- exact solutions
- delta sensitivity
- training data
- priors’ validity

1979: Decision Graph (Influence diagram) / Bayesian network definition (Howard & Matheson)
1981: First interactive universal inference algorithm solution (Shachter)
1984: Singly connected Bayesian networks inference algorithm solution (Pearl)
1991: Second universal solution – symbolic probabilistic inference algorithm (D’Ambrosio)
1992: All exact universal decision algorithm solutions proven to be NP-hard (Cooper)
1997: IET trade-secret Java implementation of SPI (Takikawa & D’Ambrosio)
1998 – present: IET Quiddity*Suite for additional functionality; object-orientation; robustness
Bayesian Network Solutions

Join Tree

Symbolic Probabilistic Inference

Message Passing Algorithm
Numerical Computation
Must be repeated for every query

Symbolical mathematics to simplify the equation
Then compute numerical result
Unchanged factors can be cached
Each query is a function
Can generate extremely efficient runtime code

\[ P(D, E) = \sum_{A} \sum_{B} \sum_{C} P(A)P(B \mid A)P(C \mid A)P(D \mid B, C)P(E \mid C) \]

• Proceedings Uncertainty in Artificial Intelligence, AUAI Press
• www.auai.org
Naive Evidential Accrual

Hypothesis:

Evidence:

Prior Conditionals:
Naive Evidential Accrual - Prior

Netica - www.norsys.com

<table>
<thead>
<tr>
<th>Attack</th>
<th>true</th>
<th>false</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.10</td>
<td>99.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Telephone_Conversation</th>
<th>true</th>
<th>false</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>false</td>
<td>50.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meeting_Observed</th>
<th>true</th>
<th>false</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>false</td>
<td>96.9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repeated_Presence_Observed</th>
<th>true</th>
<th>false</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>false</td>
<td>95.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Purchase_Record</th>
<th>true</th>
<th>false</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>35.0</td>
<td></td>
</tr>
<tr>
<td>false</td>
<td>65.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Safe_House_Observation</th>
<th>true</th>
<th>false</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>false</td>
<td>50.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>a:=true</th>
<th>a:=false</th>
</tr>
</thead>
<tbody>
<tr>
<td>p(t:=true</td>
<td>a)</td>
</tr>
<tr>
<td>p(t:=false</td>
<td>a)</td>
</tr>
</tbody>
</table>
Naive Evidential Accrual - Posterior
Dependent Evidential Accrual - Prior

\begin{align*}
\text{a:=true} & \quad \text{a:=false} \\
\text{m:=true} & \quad \text{m:=false}
\end{align*}

\begin{align*}
p(t:=\text{true} | m & a) & \quad .5 & \quad .7 & \quad .7 & \quad .1 \\
p(t:=\text{false} | m & a) & \quad .5 & \quad .3 & \quad .3 & \quad .9
\end{align*}
Dependent Evidential Accrual - Posterior
Causal Models Intermediate Variables

Hypothesis:

Steps:

Evidence:

Prior Conditionals: \( p(\text{step} | \text{attack}) \); \( p(\text{evidence} | \text{step}) \)

Result: Causal representation makes prior conditionals reasonably estimable
Causal Evidential Accrual - Prior

\[
p(t := \text{true} | o) = 0.7 \\
p(t := \text{false} | o) = 0.3 \\
p(o := \text{true} | a) = 0.999 \\
p(o := \text{false} | a) = 0.001 \\
p(o := \text{true} | \overline{a}) = 0.1 \\
p(o := \text{false} | \overline{a}) = 0.9
\]
Counter-Terrorism Reasoning

• Use causal Bayesian probabilistic modeling for rare event inductive reasoning
  – Causality provides natural linkages for observable evidence
  – Causality constrains the sampling space

• Develop causally structured models of variable relations that support reasoning chains
  – Causal models have “natural” conditional priors
  – Provides platform for discussion of prior probabilities

• Account for hypotheses completeness and credibility
  – Models introduce credibility and context variables
  – “Forces” consideration of complementary alternative hypotheses

• Arrival-driven automated evidence accrual
  – Account for variables dependencies
  – Solve problem of asynchronous evidence arrival