Generalized Information Theory for Hints

Marc Pouly

marc.pouly@uni.lu
Interdisciplinary Centre for Security, Reliability and Trust
University of Luxembourg

June 2011

securityandtrust.lu

Hartley's Measure (1928)

Given a set $S = \{s_1, \ldots, s_n\}$ how can we measure its uncertainty $u(S)$

1. uncertainty is a non-negative value
2. monotone: $|S_1| \leq |S_2| \Rightarrow u(S_1) \leq u(S_2)$
3. additive: $u(S_1 \times S_2) = u(S_1) + u(S_2)$

Theorem: There is only one function that satisfies these requirements

$$\boxed{u(S) \;=\; \log|S|}$$

Shannon's Measure (1948)

Given a set $S = \{s_1, \ldots, s_n\}$ with probabilities $p_i = p(s_i)$ how can we measure its uncertainty $u(S)$

$$\boxed{S(p_1, \ldots, p_n) \;=\; -\sum_{i=1}^{n} p_i \log p_i}$$

- We have similar uniqueness results for specific requirements
- Shannon generalizes Hartley: $S(\frac{1}{n}, \ldots, \frac{1}{n}) = \log n$

Uncertainty Measure for Dempster-Shafer Theory

What is the uncertainty of a mass function $\rightsquigarrow$ Requirements

1. **Generalization** of the Shannon and Hartley measure
2. **Additivity:** for non-interactive mass functions

$$GS(m_1 \otimes m_2) \quad = \quad GS(m_1) + GS(m_2)$$

3. **Subadditivity:** if m is defined over partition $s \cup t$

$$GS(m) \quad \leq \quad GS(m^{\downarrow s}) + GS(m^{\downarrow t})$$

4. **Expansibility, Symmetry, Continuity, Normalization**

The Aggregate Uncertainty

In 1994 several authors independently proposed

$$AU(m) \quad = \quad \max_{\mathcal{P}_{bel}} \left[-\sum_{\theta \in \Theta} p(\theta) \log p(\theta) \right]$$

- This measure is called Aggregate Uncertainty
- It satisfies all the requirements above
- The max is taken over all probability distributions that dominate *bel*

Shortcomings of the Aggregate Uncertainty

- Computing uncertainty requires to solve a non-linear optimization problem. Although an algorithm exists, computing the uncertainty of even simplest mass functions by hand is impossible (for me).
- It has been shown that the Aggregate Uncertainty is highly insensitive to changes in evidence.
- Anyway, it does (!) satisfy the requirements.

Shortcomings of the Aggregate Uncertainty

- Computing uncertainty requires to solve a non-linear optimization problem. Although an algorithm exists, computing the uncertainty of even simplest mass functions by hand is impossible (for me).
- It has been shown that the Aggregate Uncertainty is highly insensitive to changes in evidence.
- Anyway, it does (!) satisfy the requirements.

Uniqueness of the Aggregate Uncertainty

Citation from (Harmanec, 1999) page 6:

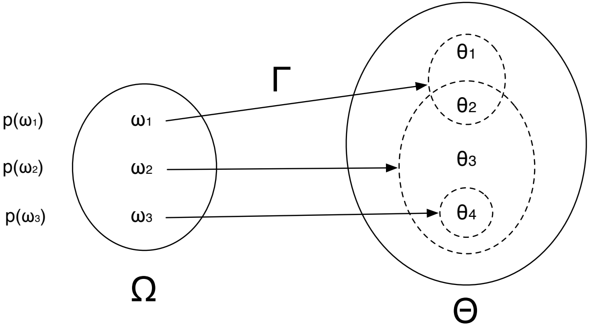
> *An interesting question is whether the measure AU is the only measure satisfying the requirements. It is still an open question ...*

Citation from (Klir, 2005) page 228:

> *It is thus a well-justified measure [...] in Dempster-Shafer Theory. Although its uniqueness is still an open problem ...*

The Theory of Hints (Kohlas & Monney, 1995)

- A particular approach to Dempster-Shafer theory
- Do hints give a new perspective on uncertainty ?



The Theory of Hints

1. A hint refers to a question with possible answers $\Theta = \{\theta_1, \dots, \theta_n\}$
2. Set of configurations (states) of the actual world $\Omega = \{\omega_1, \dots, \omega_m\}$
3. If $\omega \in \Omega$ holds, the true answer belongs to $\Gamma(\omega) \subseteq \Theta$ (causal relation)
4. Each $\omega \in \Omega$ has a probability $p(\omega)$
5. Hint: $\mathcal{H} = (\Theta, \Omega, \Gamma, p)$ where $\Gamma : \Omega \rightarrow \mathcal{P}(\Theta)$

Hints and Dempster-Shafer Theory

- Hints give information over $\Omega \times \Theta$, mass functions over Θ
- Hints induce mass functions, i.e. for $A \subseteq \Theta$ we define

$$m(A) \quad = \quad \sum_{\omega \in \Omega : \Gamma(\omega) = A} p(\omega)$$

- Hints are equivalent if they induce the same mass function
- Hints therefore provide more fine-grained information, i.e. one mass function stands for an equivalence class of hints

Hints and Dempster-Shafer Theory

- Hints give information over $\Omega \times \Theta$, mass functions over Θ
- Hints induce mass functions, i.e. for $A \subseteq \Theta$ we define

$$m(A) = \sum_{\omega \in \Omega : \Gamma(\omega) = A} p(\omega)$$

- Hints are equivalent if they induce the same mass function
- Hints therefore provide more fine-grained information, i.e. one mass function stands for an equivalence class of hints

Uncertainty of Hints

- Given $\omega \in \Omega$ the true answer is in $\Gamma(\omega)$
- There is not more evidence $\leadsto$ Hartley is the only justified measure: $H(\mathcal{H}|\omega) = \log|\Gamma(\omega)|$
- Hartley equals Shannon on uniform distributions, we set

$$p(\theta|\omega) = \frac{1}{|\Gamma(\omega)|} \quad \text{for all } \theta \in \Gamma(\omega)$$

- Therefore $p(\omega,\theta) = p(\theta|\omega) \cdot p(\omega) = \frac{p(\omega)}{|\Gamma(\omega)|}$ for $\theta \in \Gamma(\omega)$

Uncertainty of Hints

- Given $\omega \in \Omega$ the true answer is in $\Gamma(\omega)$
- There is not more evidence $\rightsquigarrow$ Hartley is the only justified measure: $H(\mathcal{H}|\omega) = \log|\Gamma(\omega)|$
- Hartley equals Shannon on uniform distributions, we set

$$p(\theta|\omega) = \frac{1}{|\Gamma(\omega)|} \quad \text{for all } \theta \in \Gamma(\omega)$$

- Therefore $p(\omega, \theta) = p(\theta|\omega) \cdot p(\omega) = \frac{p(\omega)}{|\Gamma(\omega)|}$ for $\theta \in \Gamma(\omega)$

The Hints Entropy

We measure the uncertainty of a hint $\mathcal{H} = (\Theta, \Omega, \Gamma, p)$ by Shannon's entropy applied to the joint distribution $p(\omega, \theta)$:

$$\boxed{H(\mathcal{H}) = -\sum_{\omega \in \Omega} \sum_{\theta \in \Gamma(\omega)} p(\omega, \theta) \log p(\omega, \theta)}$$

Track Record of the Hints Entropy

1. **Generalization** of the Shannon and Hartley measure $\checkmark$
2. **Additivity:** for non-interactive hints

$$H(\mathcal{H}_1 \otimes \mathcal{H}_2) \quad = \quad H(\mathcal{H}_1) + H(\mathcal{H}_2) \;\checkmark$$

3. **Subadditivity:** if $\mathcal{H}$ is defined over partition $s \cup t$

$$H(\mathcal{H}) \quad \leq \quad H(\mathcal{H}^{\downarrow s}) + H(\mathcal{H}^{\downarrow t}) \;\checkmark$$

4. **Expansibility, Symmetry, Continuity, Normalization** $\checkmark$

Hints Entropy in the Literature

How is the hints entropy related to other functionals ?

- From $p(\omega,\theta) = \frac{p(\omega)}{|\Gamma(\omega)|}$ for $\theta \in \Gamma(\omega)$ follows

$$H(\mathcal{H}) = -\sum_{\omega\in\Omega} p(\omega) \log p(\omega) + \sum_{\omega\in\Omega} p(\omega) \log |\Gamma(\omega)|$$

- 2nd term is known as the Generalized Hartley Measure

- The hints entropy is the sum of two well-studied functionals

Hints Entropy in the Literature

How is the hints entropy related to other functionals ?

- From $p(\omega, \theta) = \frac{p(\omega)}{|\Gamma(\omega)|}$ for $\theta \in \Gamma(\omega)$ follows

$$H(\mathcal{H}) = -\sum_{\omega \in \Omega} p(\omega) \log p(\omega) + \sum_{\omega \in \Omega} p(\omega) \log |\Gamma(\omega)|$$

- 2nd term is known as the Generalized Hartley Measure
- The hints entropy is the sum of two well-studied functionals

Discussion

The hints entropy ...

- ... satisfies the same properties as the Aggregate Uncertainty
- ... does not suffer from the sensitivity problem
- ... is defined as an explicit formula, which is easy to compute

Does this disprove uniqueness of the Aggregate Uncertainty ?

No, because the two functionals have different range

- Aggregate Uncertainty: $0 \leq AU(m) \leq \log|\Theta|$
- Hints Entropy: $0 \leq H(m) \leq \log|\Omega \times \Theta|$

Discussion

The hints entropy ...

... satisfies the same properties as the Aggregate Uncertainty

... does not suffer from the sensitivity problem

... is defined as an explicit formula, which is easy to compute

Does this disprove uniqueness of the Aggregate Uncertainty ?

No, because the two functionals have different range

- Aggregate Uncertainty: $0 \leq AU(m) \leq \log |\Theta|$
- Hints Entropy: $0 \leq H(m) \leq \log |\Omega \times \Theta|$

Pignistic Probabilities

- Hints entropy defined on $\Omega \times \Theta \rightsquigarrow$ marginalize to Θ
- We had $p(\omega, \theta) = \frac{p(\omega)}{|\Gamma(\omega)|}$ for $\theta \in \Gamma(\omega)$
- The marginal distribution therefore is

$$p(\theta) = \sum_{\omega \in \Omega} p(\omega, \theta) = \sum_{\omega \in \Omega : \theta \in \Gamma(\omega)} \frac{p(\omega)}{|\Gamma(\omega)|}$$

- This is called the pignistic distribution of the hint
- Equivalent hints share the same pignistic distribution

Pignistic Probabilities

- Hints entropy defined on $\Omega \times \Theta \rightsquigarrow$ marginalize to Θ
- We had $p(\omega, \theta) = \frac{p(\omega)}{|\Gamma(\omega)|}$ for $\theta \in \Gamma(\omega)$
- The marginal distribution therefore is

$$p(\theta) = \sum_{\omega \in \Omega} p(\omega, \theta) = \sum_{\omega \in \Omega : \theta \in \Gamma(\omega)} \frac{p(\omega)}{|\Gamma(\omega)|}$$

- This is called the pignistic distribution of the hint
- Equivalent hints share the same pignistic distribution

Pignistic Entropy

- The hints entropy restricted to Θ is the pignistic entropy

$$\boxed{H(\Theta) = -\sum_{\theta \in \Theta} p(\theta) \log p(\theta)}$$

where

$$p(\theta) = \sum_{\omega \in \Omega : \theta \in \Gamma(\omega)} \frac{p(\omega)}{|\Gamma(\omega)|}$$

- Equivalent hints have the same pignistic entropy
- Uncertainty measure for Dempster-Shafer theory

History of the Pignistic Entropy

- In 2006, the pignistic entropy was proposed as an uncertainty measure that satisfies all requirements
- In 2008, a mistake in the proof of subadditivity was pointed out and a counter-example was given
- Flaw: the pignistic distribution of a marginal $m^{\downarrow s}$ is not necessarily the marginal of the pignistic distribution of m

Conclusions

- The hints entropy on $\Omega \times \Theta$ satisfies all properties, the pignistic entropy on Θ violates subadditivity
- But the hints entropy satisfies a weaker form of subadditivity $\rightsquigarrow$ see paper
- DS theory studies equivalence classes of hints
- By looking at equivalence classes we naturally loose information $\rightsquigarrow$ destroys strong subadditivity

Late Credits

- The paper from 2006 on the pignistic entropy is flawed - the pignistic entropy does not satisfy subadditivity
- Eliminating the mistake still proves weak subadditivity
- However, the same ideas applied to hints instead of mass functions gives an uncertainty measure that satisfies all properties including strong subadditivity.


?
R
D