Efficient Computation of Belief Theoretic Conditionals

Lalintha G. Polpitiya, Kamal Premaratne, Manohar N. Murthi & Dilip Sarkar

University of Miami
Coral Gables, Florida, USA
`lalintha@umiami.edu,{kamal,mmurthi,sarkar}@miami.edu`

Challenges and Contributions

- In spite of the greater expressiveness and flexibility they offer, **Dempster-Shafer (DS) theoretic** implementations in current use are restricted to smaller frames of discernment (FoDs) because of the prohibitive computational burden that larger FoDs impose on existing methods.
- While this difficulty has been addressed via several approximation methods[1], such approaches usually require one to compromise the quality of the generated results for computational efficiency.
- Exact (or sufficiently precise) computation of conditionals is of paramount importance because the quality of results generated from DS theoretic (DST) strategies depend directly on the precision of the conditional.
- The main contribution of this paper is a completely new generalized model **(DS-Conditional-One)** for computing DST conditionals.
- This model can be employed to compute both the FH and Dempster's conditional beliefs of an **arbitrary** proposition. This is exactly the challenge that Shafer refers to in [5, p.348], viz., **"It remains to be seen how useful the fast Möbius transform will be in practice. It is clear, however, that it is not enough to make arbitrary belief function computations feasible."**
- Our experiment results demonstrate that the average computational time taken to compute the conditional belief of an arbitrary proposition by the proposed approach is less than **2 (μs) for a FoD of size 10** and **0.7 (ms) for a FoD of size 20 ($\sim$1 million focal elements).**
- This new model can also be utilized as a **visualization tool** for conditional computations and in analyzing characteristics of conditioning operations.
- An outcome of this research is a **conditional computation library** which is available at **ProFuSELab** (Scan the QR code above).

(`https://profuselab.github.io/Conditional-Computation-Library/`)

Mathematical Foundation

- The following notation is useful for our work:

$$\mathcal{S}(A;B) = \sum_{\substack{\emptyset \neq C \subseteq A;\\ \emptyset \neq D \subseteq B}} m(C \cup D). \qquad (1)$$

$\mathcal{S}(A;B)$ denotes the sum of all masses of propositions that 'straddle' both $A \subseteq \Theta$ and $B \subseteq \Theta$.

- The following result is of critical importance for our work:

Proposition 1 *Consider the body of evidence (BoE) $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \subseteq \Theta$. For $B \subseteq \Theta$, consider the mappings $\Gamma_A : 2^{\Theta} \mapsto [0,1]$ and $\Pi_A : 2^{\Theta} \mapsto [0,1]$, where*

$$\Gamma_A(B) = \sum_{\emptyset \neq X \subseteq \overline{A}} m((A \cap B) \cup X); \quad \Pi_A(B) = \sum_{Y \subseteq (A \cap B)} \Gamma_A(Y).$$

Then the following are true:
***(i)** $\Gamma_A(A \cap B) = \Gamma_A(B)$ and $\Pi_A(A \cap B) = \Pi_A(B)$. So, w.l.o.g., we assume that $B \subseteq A$.*
***(ii)** $\Gamma_A(\emptyset) = Bl(\overline{A})$.*

- Fagin-Halpern (FH) conditional can be considered the most natural generalization of the probabilistic conditional notion because of its close connection with the inner and outer conditional probability measures in probability theory.

Definition 2 (Fagin-Halpern (FH) Conditional) *[2] Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \in \widehat{\mathfrak{F}}$. The conditional belief $Bl(B|A)$ of B given the conditioning event A is*

$$Bl(B|A) = \frac{Bl(A \cap B)}{Bl(A \cap B) + Pl(A \cap \overline{B})}. \qquad \blacksquare$$

- For our work, we need the following alternate expression for the FH conditional:

Proposition 3 *Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \in \widehat{\mathfrak{F}}$. Then, we may express $Bl(B|A)$ as*

$$Bl(B|A) = \frac{Bl(A \cap B)}{1 - Bl(\overline{A}) - \mathcal{S}(\overline{A}; A \cap B)}, \; B \subseteq \Theta. \qquad \square$$

FH conditioning annuls those propositions that 'straddle' the conditioning proposition A and its complement $\overline{A}$.

- Dempster's conditional is perhaps the most widely employed DST conditional notion.

Definition 4 (Dempster's Conditional) *[4] Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \subseteq \Theta$ s.t. $Bl(\overline{A}) \neq 1$, or equivalently, $Pl(A) \neq 0$. The conditional belief $Bl(B\|A)$ of B given the conditioning event A is*

$$Bl(B\|A) = \frac{Bl(\overline{A} \cup B) - Bl(\overline{A})}{1 - Bl(\overline{A})}. \qquad \blacksquare$$

Dempster's conditioning also annuls masses of all those propositions that 'straddle' the conditioning proposition A and its complement $\overline{A}$.

- For our work, we need the following alternate expression for the Dempster's conditional:

Proposition 5 *Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \subseteq \Theta$ s.t. $Bl(\overline{A}) \neq 1$. Then, $Bl(B\|A)$ can be expressed as*

$$Bl(B\|A) = \frac{Bl(A \cap B) + \mathcal{S}(\overline{A}; A \cap B)}{1 - Bl(\overline{A})}, \; B \subseteq \Theta. \qquad \square$$

- **REGAP** (**RE**cursive **G**eneration of and **A**ccess to **P**ropositions) property[3]:

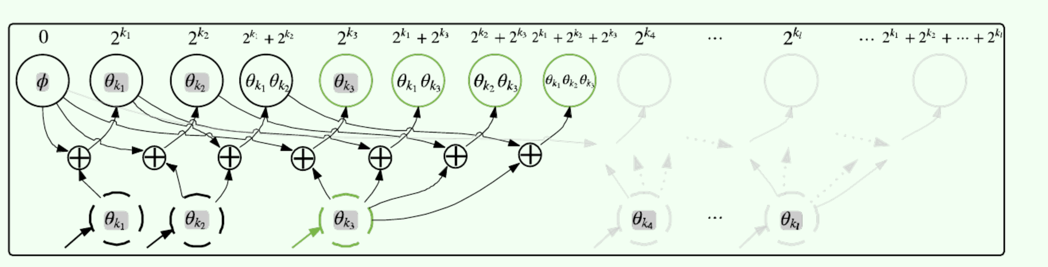

Fig. 1: REGAP: REcursive Generation of and Access to Propositions.

DS-Conditional-One Computational Model

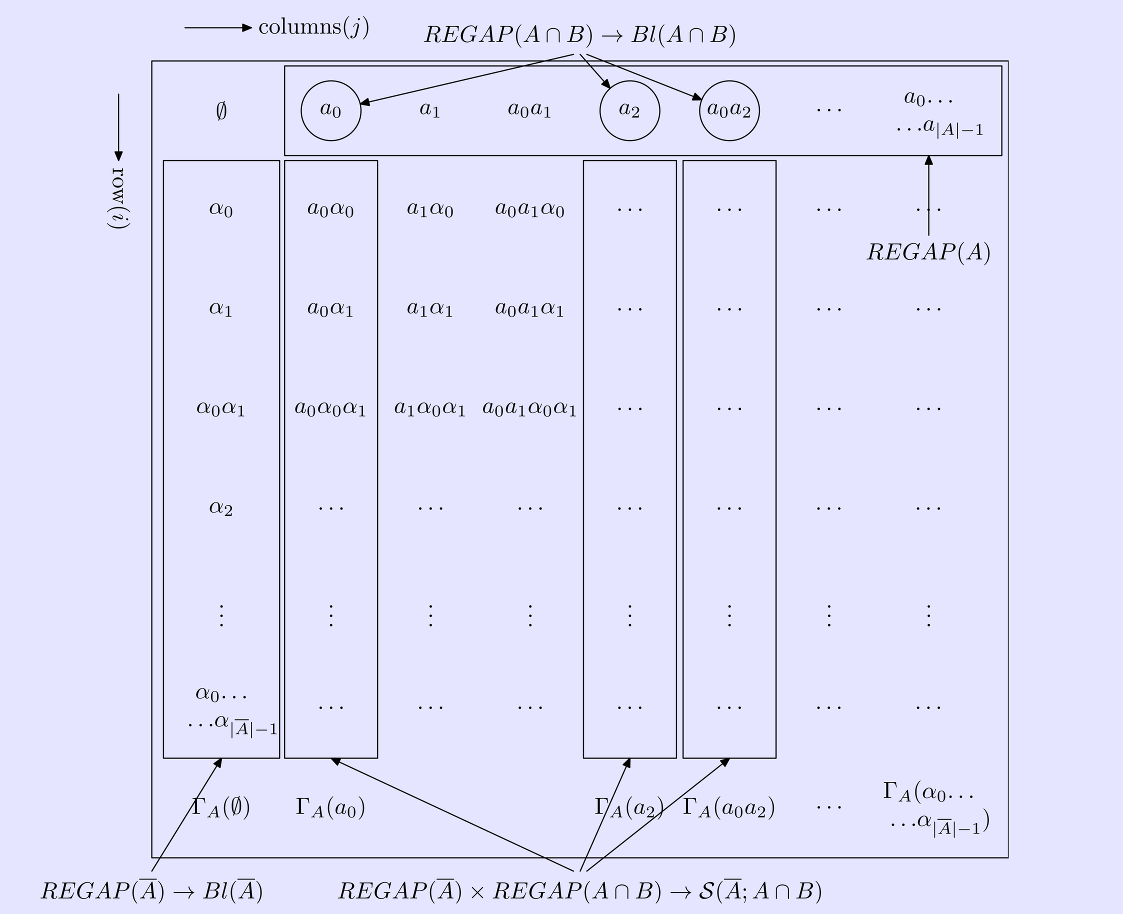

Fig. 2: DS-Conditional-One model. Quantities related to $Bl(B|A)$ computation when $A = \{a_0, a_1, \ldots, a_{|A|-1}\}$ and $\overline{A} = \{\alpha_0, \alpha_1, \ldots, \alpha_{|\overline{A}|-1}\}$, and $B = \{a_0, a_2\} \subseteq A$.

- **DS-Conditional-One** is a computational model that enables one to compute the FH and Dempster's conditional beliefs of an **arbitrary** proposition.
- We denote the conditioning proposition A, its complement $\overline{A}$, and the conditioned proposition B as $\{a_0, a_1, \ldots, a_{|A|-1}\}$, $\{\alpha_0, \alpha_1, \ldots, \alpha_{|\overline{A}|-1}\}$, and $\{b_0, b_1, \ldots, b_{|B|-1}\}$, respectively. Here, $\Theta = \{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$ denotes the FoD and $a_i, \alpha_j, b_k \in \Theta$.
- We represent singletons of the conditioning event $A = \{a_0, a_1, \ldots, a_{|A|-1}\}$ as *column singletons* and singletons of the complement of conditioning event $\overline{A} = \{\alpha_0, \alpha_1, \ldots, \alpha_{|\overline{A}|-1}\}$ as *row singletons* in a DS-Matrix. See Fig. 2.
- The proposed DS-Conditional-One computational model allows direct identification of $REGAP(\overline{A})$, $REGAP(A \cap B)$, $(REGAP(\overline{A}) \times REGAP(A \cap B))$, $(REGAP(A) \times REGAP(\overline{A}))$, and $\Gamma_A(C)$, $\forall C \subseteq B$.

Compute $Bl(A \cap B)$, $Bl(\overline{A})$, and $\mathcal{S}(\overline{A}; A \cap B)$

- We use a lookup table named *power* to enhance the computational efficiency. It contains 2 to the power of singleton indexes in increasing order.
- *index*[] is a dynamic array which keeps the indexes of subset propositions of $A \cap B$.

(a) $REGAP(A \cap B)$: Use this to compute $Bl(A \cap B)$.

Algorithm 1 Compute $Bl(A \cap B)$ (with time complexity $\mathcal{O}(2^{|A \cap B|})$)

```
procedure BLB(Singletons A, Singletons B, DS-Matrix BBA)
    belief ← 0
    count ← 0
    for each a_i in A ∩ B do
        index[count] ← power[i]
        temp ← count
        count ← count + 1
        for j ← 0, temp − 1 do
            index[count] ← index[j] + power[i]
            count ← count + 1
        end for
    end for
    for i ← 0, power[|A ∩ B|] − 2 do
        belief ← belief + BBA[0][index[i]]
    end for
    Return belief
end procedure
```

(b) $REGAP(\overline{A})$: Use this to compute $Bl(\overline{A})$.

Algorithm 2 Compute $Bl(\overline{A})$ (with time complexity $\mathcal{O}(2^{|\overline{A}|})$)

```
procedure BLCOMP(Singletons Ā, DS-Matrix BBA)
    belief ← 0
    for i ← 1, power[|Ā|] − 1 do
        belief ← belief + BBA[i][0]
    end for
    Return belief
end procedure
```

(c) $(REGAP(\overline{A}) \times REGAP(A \cap B))$, the Cartesian product of $REGAP(\overline{A})$ and $REGAP(A \cap B)$: Use this to compute $\mathcal{S}(\overline{A}; A \cap B)$.

Algorithm 3 Compute $\mathcal{S}(\overline{A}; A \cap B)$ (with time complexity $\mathcal{O}(2^{|\overline{A}|+|A \cap B|})$)

```
procedure STRAD(Singletons Ā, Singletons A, Singletons B, DS-Matrix BBA)
    belief ← 0
    count ← 0
    for each a_i in A ∩ B do
        index[count] ← power[i]
        temp ← count
        count ← count + 1
        for j ← 0, temp − 1 do
            index[count] ← index[j] + power[i]
            count ← count + 1
        end for
    end for
    for i ← 1, power[|Ā|] − 1 do
        for j ← 0, power[|A ∩ B|] − 2 do
            belief ← belief + BBA[i][index[j]]
        end for
    end for
    Return belief
end procedure
```

Space Complexity of Algorithms 1, 2, and 3. The matrix in Fig. 2 is of size $2^{|A|} \times 2^{|\overline{A}|}$. Hence, the space complexity associated with each algorithm above is $\mathcal{O}(2^{|\Theta|})$.

Computation of Conditionals

- **FH Conditional Belief of an Arbitrary Proposition:** Use the expression in Proposition 3, where $Bl(A \cap B)$, $Bl(\overline{A})$ and $\mathcal{S}(\overline{A}; A \cap B)$ are obtained via Algorithms 1, 2, and 3, respectively.

The computational complexity is $\mathcal{O}(2^{|\overline{A}|+|A \cap B|})$.
$Bl(B|A)$ for $B = \{a_0, a_2\}$ is computed as (See Fig. 2)

$$Bl(B|A) = \frac{Bl(A \cap B)}{1 - \Gamma_A(\{\emptyset\}) - \Gamma_A(\{a_0\}) - \Gamma_A(\{a_2\}) - \Gamma_A(\{a_0, a_2\})}. \qquad (2)$$

- **Dempster's Conditional Belief of an Arbitrary Proposition:** Use the expression in Proposition 5, where $Bl(A \cap B)$, $Bl(\overline{A})$ and $\mathcal{S}(\overline{A}; A \cap B)$ are obtained via Algorithms 1, 2, and 3, respectively.

The computational complexity is $\mathcal{O}(2^{|\overline{A}|+|A \cap B|})$.
Consider the same example as before, viz., $B = \{a_0, a_2\}$. Then, we may compute $Bl(B\|A)$ as

$$Bl(B\|A) = \frac{Bl(A \cap B) + \Gamma_A(\{a_0\}) + \Gamma_A(\{a_2\}) + \Gamma_A(\{a_0, a_2\})}{1 - \Gamma_A(\{\emptyset\})}. \qquad (3)$$

- **Dempster's Conditional Mass Using Specialization Matrix[6]:**

It employs a $2^{|\Theta|} \times 2^{|\Theta|}$-sized stochastic matrix $\mathfrak{S}_A$ (with each entry '0' or '1') referred to as the conditioning specialization matrix and a $2^{|\Theta|} \times 1$-sized vector $m(\cdot)$ containing the focal elements. The computational and space complexity of the specialization matrix multiplication is $\mathcal{O}(2^{|\Theta|} \times 2^{|\Theta|})$, a prohibitive burden even for modest FoD sizes.

Experiments and Concluding Remarks

- For a given FoD size, we selected a random set of focal elements, with randomly selected mass values, and conducted 10,000 conditional computations for randomly chosen propositions A and $B \subseteq A$.
- With the DS-Conditional-One model (which applies to *both* FH and Dempster's conditionals), we use a 'brute force' approach to compute all the conditional beliefs. We then use the fast Möbius transform (FMT) [7] to get the conditional masses for all the propositions.
- All conditional computations for an **arbitrary** proposition were done on an iMac running Mac OS X 10.12.3 (with 2.9GHz Intel Core i5 processor and 8GB of 1600MHz DDR3 RAM). Conditional computations for **all** propositions were done on the same iMac for smaller FoDs and on a supercomputer (`https://ccs.miami.edu/pegasus`) for larger FoDs (underlined in Table 1).

Method → Conditional →		DS-Conditional-One Model FH or Dempster's			Specializa. Dempster's
FoD $\|\Theta\|$	Max. $\|\mathfrak{F}\|$	$Bl(B\|A)$ or $Bl(B\|\|A)$ (Arbitrary)	$Bl(B\|A)$ or $Bl(B\|\|A)$ (All)	$m(B\|A)$ or $m(B\|\|A)$ (All)	$m(B\|\|A)$ (All)
2	3	0.0005	0.0011	0.0016	0.0011
4	15	0.0005	0.0038	0.0050	0.0063
6	63	0.0006	0.0128	0.0170	0.0696
8	255	0.0009	0.0517	0.0679	1.0154
10	1,023	0.0017	0.2428	0.3090	93.1590
12	4,095	0.0040	1.3528	1.6186	1,485.6300
14	16,383	0.0120	18.4885	22.4995	25,051.8200
16	65,535	0.0405	146.1480	151.9600	***
18	262,143	0.1516	1,087.2800	1,113.5300	***
20	1,048,575	0.6011	8,485.4500	8,862.9800	***

Table 1: DS-Conditional-One model versus specialization matrix based method. Average computational times (ms). (*** denotes computations not completed within a feasible time).

- The significant speed advantage offered by the proposed computational model over the specialization matrix based approach is evident from Table 1.
- For larger FoDs, the computational burden associated with the specialization matrix based approach becomes prohibitive because of its space complexity of $\mathcal{O}(2^{|\Theta|} \times 2^{|\Theta|})$. For example, an FoD of size 20 would need 128 ($= 2^{20} \times 2^{20}/8$) GB of memory to represent the specialization matrix, if each matrix entry occupies only 1 bit.
- Another advantage of the proposed approach is that it can be utilized for either the FH conditional or Dempster's conditional belief computations.
- It also appears possible to further enhance the algorithms that we have developed via parallel computing optimizations because of the underlying matrix structure.

Acknowledgments

This work is based on research supported by the U.S. Office of Naval Research (ONR) via grant #N00014-10-1-0140 and the U.S. National Science Foundation (NSF) via grant #1343430.

References

[1] T. Denœux. 40 years of Dempster–Shafer theory. *Int. J. Approx. Reason.*, 79(C):1–6, Dec 2016.

[2] R. Fagin and J. Y. Halpern. A new approach to updating beliefs. In *Proc. 6th Conf. Uncertainty in Artificial Intelligence (UAI)*, pages 347–374, Cambridge, MA, Jul 1990.

[3] L. G. Polpitiya, K. Premaratne, M. N. Murthi, and D. Sarkar. A framework for efficient computation of belief theoretic operations. In *Proc. 19th Int. Conf. Information Fusion (FUSION)*, pages 1570–1577, Heidelberg, Germany, Jul 2016.

[4] G. Shafer. *A Mathematical Theory of Evidence.* Princeton Univ. Press, Princeton, NJ, 1976.

[5] G. Shafer. Perspectives on the theory and practice of belief functions. *Int. J. Approx. Reason.*, 4(5-6):323–362, Oct 1990.

[6] P. Smets. The application of the matrix calculus to belief functions. *Int. J. Approx. Reason.*, 31(1):1–30, 2002.

[7] H. M. Thoma. *Factorization of belief functions.* PhD thesis, Dept. Stat., Harvard Univ., Cambridge, MA, 1989.