PHD THESIS

METHODS OF PROBABILISTIC KNOWLEDGE INTEGRATION

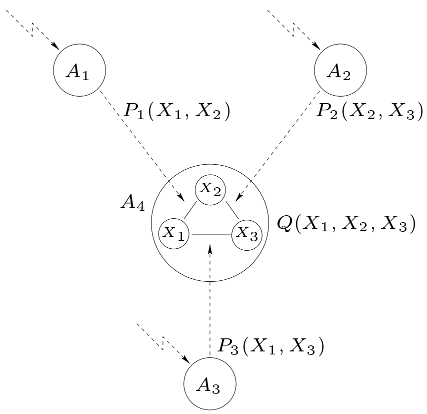
JIŘÍ VOMLEL

vomlel@lisp.vse.cz

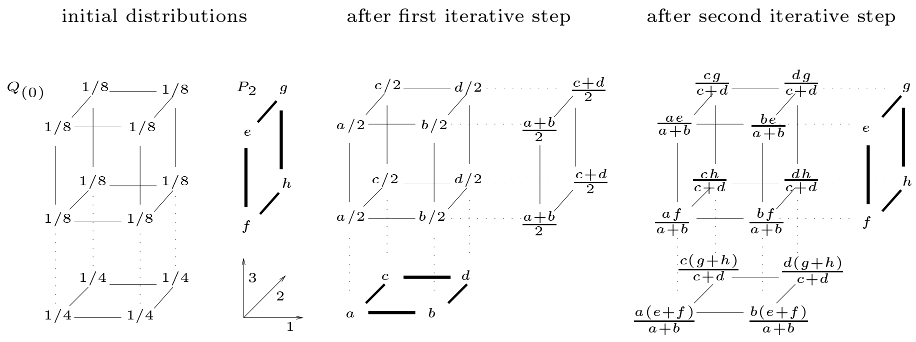
PRAGUE, SEPTEMBER 21, 2000

A KNOWLEDGE INTEGRATION STORY

- There are three human experts with narrow knowledge of their domains.
- These human experts are to be replaced by software agents A_1, A_2, and A_3.
- Software agents are trained to be as good as the domain experts.
- The knowledge of each agent is represented by a few-dimensional probability distribution, $P_1(X_1, X_2)$, $P_2(X_2, X_3)$, and $P_3(X_1, X_3)$.
- The human experts got out of reach.
- A procedure should be designed so that an software agent A_4 is able to *integrate knowledge* of agents A_1, A_2, and A_3.
- Furthermore, agent A_4 should be able to update its model if agents A_1, A_2, and A_3 update their models in light of a new evidence.



ITERATIVE PROPORTIONAL FITTING PROCEDURE (IPFP)



$Q_{(1)} = Q_{(0)} \frac{P_1}{Q_{(0)}^{\{1,2\}}}$ $\quad$ $Q_{(2)} = Q_{(1)} \frac{P_2}{Q_{(1)}^{\{2,3\}}}$

If $a+b = e+f$ and $c+d = g+h$ (weak consistency of P_1 and P_2) then $Q_{(2)}^{\{1,2\}} = \{a,b,c,d\} = P_1$

BASIC DEFINITIONS AND TWO BASIC PROPERTIES OF IPFP

$\mathcal{P}$	set of all probability distributions defined for $\{X_i\}_{i \in V}$
$P \ll Q$	P is dominated by Q iff $\{x, P(x) > 0\} \subseteq \{y, Q(y) > 0\}$
$\mathcal{E}$	*generating class* of input set with an ordering $\{E_1, E_2, \dots E_s\}$
$\mathcal{S}_{\mathcal{E}}$	*additively constrained set*, $\{Q \in \mathcal{P}, Q^{E_1} = P_{E_1}, \dots, Q^{E_s} = P_{E_s}\}$
$\mathcal{R}_{\mathcal{E}}$	*multiplicatively constrained set*, $\{Q \in \mathcal{P}, Q = \prod_{i=1}^{s} \psi_{E_i}(x^{E_i}),\ \psi_{E_i} : \mathbb{X}^{E_i} \mapsto \mathbb{R}, i = 1, 2, \dots, s\}$

$$I(P \parallel Q) = \sum_{x \in \mathbb{X}} P(x) \log \frac{P(x)}{Q(x)}, \text{ Kullback-Leibler divergence}$$

$$\pi_{\mathcal{S}_{\mathcal{E}}} Q = \arg\min_{\{P \in \mathcal{S}_{\mathcal{E}}\}} I(P \parallel Q), \ I_1\text{-projection to } \mathcal{S}_{\mathcal{E}}$$

IPFP properties:

- $Q_{(i)} = \boxed{\pi_{\mathcal{S}_{\{E_j\}}} Q_{(i-1)}}$ (Ireland, Kullback, 1968)
- If $\exists Q \in \mathcal{S}_{\mathcal{E}}$, $Q \ll Q_{(0)}$ then IPFP converges to $Q^* = \pi_{\mathcal{S}_{\mathcal{E}}} Q_{(0)}$. (Csiszár, 1975)

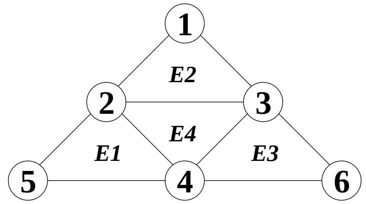
NEW CONDITION ON STRUCTURE OF $\mathcal{E}$ FOR CONVERGENCE WITHIN ONE CYCLE

(A1) $Q_{(0)} \in \mathcal{R}_{\mathcal{E}}$,

(A2) $\exists Q \in \mathcal{S}_{\mathcal{E}}$ such that $Q \ll Q_{(0)}$, and

(A3) $\exists E^* \in \mathcal{E}$ such that $E_i \cap E_j \subset E^*, \forall i, j = 1, 2, \dots s, i \neq j$

Then IPFP converges within s steps (one cycle) no matter how $\mathcal{E}$ is ordered.



$E^* = E_4$

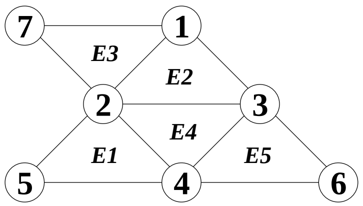
NEW CONDITION ON ORDERING OF $\mathcal{E}$ FOR CONVERGENCE WITHIN ONE CYCLE

(A1) $Q_{(0)} \in \mathcal{R}_{\mathcal{E}}$,

(A2) $\exists Q \in \mathcal{S}_{\mathcal{E}}$ such that $Q \ll Q_{(0)}$, and

(B3) for every iteration $1 < i \leq s$ of IPFP the system $\mathcal{E}_{(i)}$ iteratively constructed by $\mathcal{E}_{(i)} = \{E_1, E_2, \dots, E_i\}$ is decomposable

Then IPFP converges within s steps (one cycle).

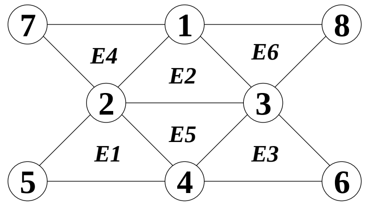


$$\mathcal{E}_{(1)} = \{E_1\},\ \mathcal{E}_{(2)} = \{E_1, E_2\},\ \mathcal{E}_{(3)} = \{E_1, E_2, E_3\},$$
$$\mathcal{E}_{(4)} = \{E_1, E_2, E_3, E_4\},\ \mathcal{E}_{(5)} = \{E_1, E_2, E_3, E_4, E_5\}$$

SHELBY HABERMAN'S CONJECTURE

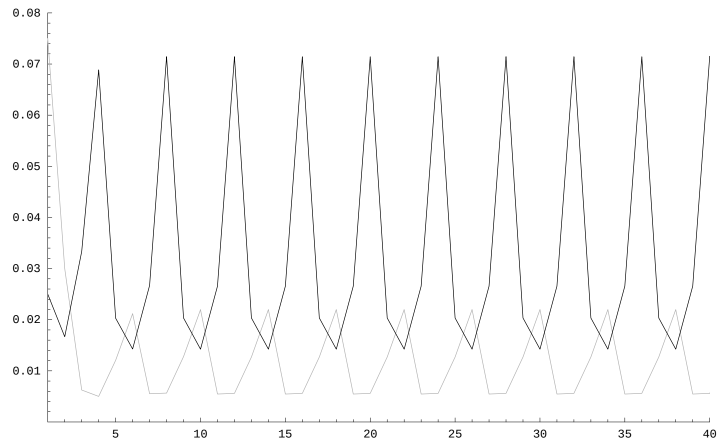
Shelby Haberman (1974) in the book *The Analysis of Frequency Data* (page 198): *"Failure to converge in one cycle does not preclude convergence in two cycles. In fact, no decomposable class is known for which convergence does not occur within two cycles."*

Counterexample:



If the ordering of $\mathcal{E}$ is $\{E_1, E_2, E_3, E_4, E_5, E_6\}$ then IPFP need not converge in two cycles.

IPFP on inconsistent input - example



upper black line - ordering $\{E_1, E_2, E_3, E_4\}$
lower grey line - ordering $\{E_4, E_2, E_1, E_3\}$

IPFP on inconsistent input - summary

I1 Convergence	**no**		
	limit cycle $\{Q_1^*, \ldots, Q_s^*\}$	arith. aver. of limit cyc. Q_a^*	geom. aver. of limit cyc. Q_g^*
I2 Ordering of $\mathcal{P}_\mathcal{E}$ does not matter	*no*	*no*	*no*
I3 I_1-aggregate minimizer $\sum_{j=1}^s w_j \cdot I(P_{E_j} \parallel Q^{E_j})$	**no**	*no*	**no**
I4 I_2-aggregate minimizer $\sum_{j=1}^s w_j \cdot I(Q^{E_j} \parallel P_{E_j})$	**no**	**no**	**no**

Methods for inconsistent input - CC and LCC

- $\{\alpha_i\}_{i=1}^\infty$ - *monotonous decreasing sequence* such that
  for $i = 1, 2, \ldots$: $\quad 0 \le \alpha_i < 1 \quad$ and $\quad \sum_{n=i}^\infty \alpha_n = +\infty$,
  $\{\alpha_i\}$ is stable within one cycle, and $\quad \lim_{i \to \infty} \alpha_i = 0$.
- Recall, that i step of regular IPFP corresponds to

$$\boxed{\pi_{\mathcal{S}_{\{E_j\}}} Q_{(i-1)}}, \quad j = ((i-1) \bmod s + 1)$$

Computational procedures:

- **CC**, conservative modif. of IPFP with convex mixture:

$$Q_{(i)} \;=\; (1-\alpha_i) \cdot Q_{(i-1)} \;+\; \alpha_i \cdot \boxed{\pi_{\mathcal{S}_{\{E_j\}}} Q_{(i-1)}}$$

- **LCC**, conser. modif. of IPFP with log-convex mixture:

$$Q_{(i)} \;=\; \frac{1}{c} \cdot Q_{(i-1)}^{1-\alpha_i} \cdot \boxed{\pi_{\mathcal{S}_{\{E_j\}}} Q_{(i-1)}}^{\alpha_i}$$

Methods for inconsistent input - AA and GA

- w_j - nonnegative weights assigned to every P_{E_j}: $\sum_{j=1}^{s} w_j = 1$.

Computational procedures:

- **AA**, method of iterative arithmetic averages:

$$Q_{(i)} \;=\; \sum_{j=1}^{s} w_j \cdot \boxed{\pi_{\mathcal{S}_{\{E_j\}}} Q_{(i-1)}}$$

- **GA**, method of iterative geometric averages:

$$Q_{(i)} \;=\; \frac{1}{c} \cdot \prod_{j=1}^{s} \boxed{\pi_{\mathcal{S}_{\{E_j\}}} Q_{(i-1)}}^{w_j}$$

Methods for inconsistent input - GEM

Two classes of subsets of V:

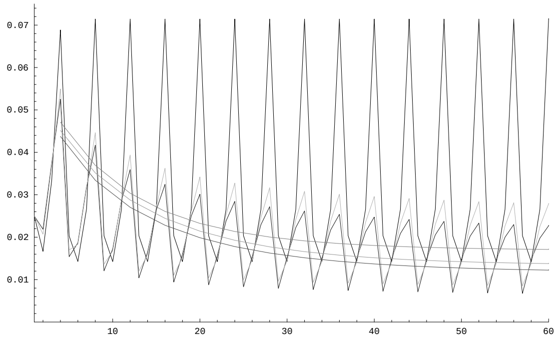
- input structure is defined by $\mathcal{E} = \{E_1, \ldots, E_s\}$
- model structure is defined by $\mathcal{F} = \{F_1, \ldots, F_t\}$

Generally, $\mathcal{E} \neq \mathcal{F}$.

An instance of generalized EM-algorithm - **GEM**:

$$
\begin{array}{llll}
\textbf{E-step:} & R_{(i)}^{F_j} & = & \left(\sum_{k=1}^{s} w_k \cdot \boxed{\pi_{\mathcal{S}_{\{E_k\}}} Q_{(i-1)}}\right)^{F_j}, \quad j = 1, 2, \ldots, t \\
\text{define} & Q_{(i,0)} & = & Q_{(i-1)} \\
\text{define} & \mathcal{S}'_{\{F_j\}} & = & \{P \in \mathcal{P} : P^{F_j} = R_{(i)}^{F_j}\} \\
\textbf{M-step:} & Q_{(i,j)} & = & \boxed{\pi_{\mathcal{S}'_{\{F_j\}}} Q_{(i,j-1)}}, \quad j = 1, 2, \ldots, t \\
\text{define} & Q_{(i)} & = & Q_{(i,t)}
\end{array}
$$

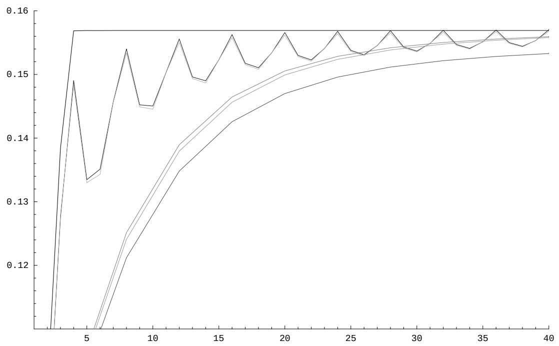
Convergence of methods on inconsistent input



The methods are (the uppermost first):

- IPFP, LCC, CC (all three are oscillating),
- GA, GEM, AA (all these three are monotonically decreasing each converges to a different value)

Convergence of methods on consistent input



The methods are (the uppermost first):

- IPFP (fast convergence),
- CC, LCC (oscillating and almost identical),
- GA, GEM (both monotonically increasing and almost identical) ,
- AA (monotonically increasing but convergent to a different value)

Methods for inconsistent input - summary

Property description		IPFP	CC	LCC	AA	GA	GEM
G1	$Q_{(0)} \in \mathcal{R}_{\mathcal{E}} \;\Rightarrow\; Q_{(i)} \in \mathcal{R}_{\mathcal{E}}, i = 1, 2, \ldots$	**yes**	**no**	**yes**	**no**	**yes**	**yes**
consistent $\mathcal{P}_{\mathcal{E}}$							
C1	Convergence	**yes**	*yes*	*yes*	**yes**	**yes**	**yes**
C2	Q^* is indep. of the ordering of $\mathcal{P}_{\mathcal{E}}$	**yes**	*yes*	*yes*	**yes**	**yes**	*yes*
C3	$Q^* = \pi_{\mathcal{E}} Q_{(0)}$	**yes**	*yes*	*yes*	*no*	**yes**	*yes*
inconsistent $\mathcal{P}_{\mathcal{E}}$							
I1	Convergence	**no**	*yes*	*yes*	**yes**	**yes**	**yes**
I2	Q^* is indep. of the ordering of $\mathcal{P}_{\mathcal{E}}$		*yes*	*yes*	**yes**	**yes**	*yes*
I3	Q^* is I_1-aggregate minimizer		*yes*	*no*	**yes**	*no*	*yes*
I4	Q^* is I_1-aggregate minimizer		*no*	*yes*	*no*	**yes**	*no*

- The study of additively and multiplicatively constrained set (original proof of Theorem 2.9)
- Two theorems giving conditions of IPFP convergence within one cycle (Theorem 4.6. and 4.7)
- Counterexample to Haberman's conjecture.
- Behavior of IPFP on inconsistent data.
- Application of GEM-algorithm to the problem of inconsistent knowledge integration.
- Comparative study of five iterative algorithms for inconsistent knowledge integration.