dfki ai Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

UNIVERSITY OF LÜBECK
INSTITUTE OF INFORMATION SYSTEMS

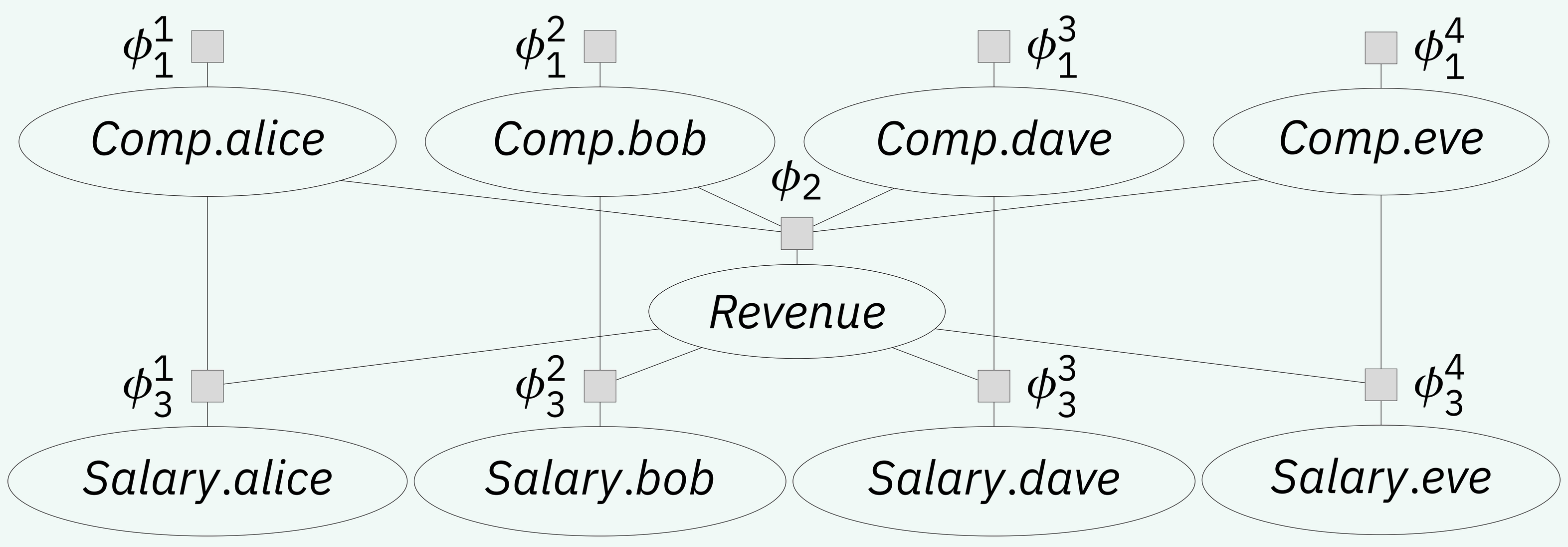
Efficient Detection of Commutative Factors in Factor Graphs

Malte Luttermann, Johann Machemer, and Marcel Gehrke

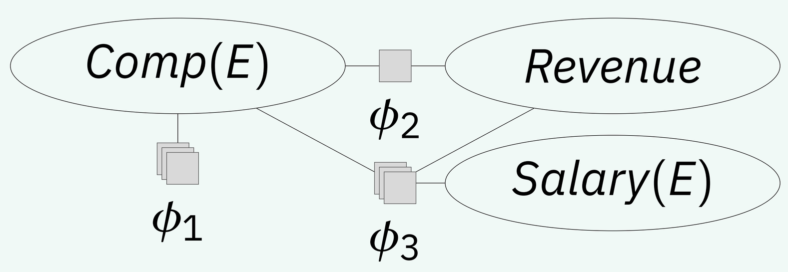
1. Motivation

- Factor graphs compactly encode a probability distribution
- Semantics of a factor graph G over a set of factors $\boldsymbol{\Phi}$:

$$P_G = \frac{1}{Z} \prod_{\phi \in \boldsymbol{\Phi}} \phi$$

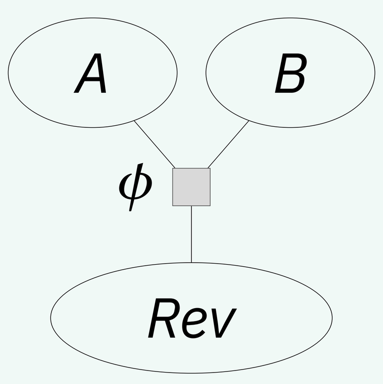


- Parametric factor graphs introduce logical variables to represent groups of random variables
- Parametric factor graphs enable lifted inference (idea: exploit indistinguishability of objects using exponentiation)



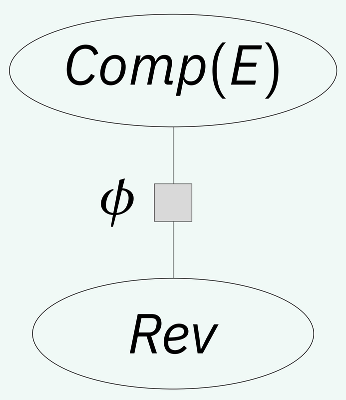
2. Problem Setup

- Goal: Efficiently detect commutative factors
- Commutative factors can be compressed



A	B	Rev	$\phi(A,B,Rev)$
true	true	true	φ_1
true	true	false	φ_4
true	false	true	φ_2
true	false	false	φ_5
false	true	true	φ_2
false	true	false	φ_5
false	false	true	φ_3
false	false	false	φ_6

↓



$\#_E[Comp(E)]$	Rev	$\phi(\#_E[Comp(E)],Rev)$
[2,0]	true	φ_1
[1,1]	true	φ_2
[0,2]	true	φ_3
[2,0]	false	φ_4
[1,1]	false	φ_5
[0,2]	false	φ_6

3. Buckets

- Buckets count occurrences of range values in an assignment

A	B	R	$\phi(A,B,R)$	b
true	true	true	φ_1	[3,0]
true	true	false	φ_4	[2,1]
true	false	true	φ_2	[2,1]
true	false	false	φ_5	[1,2]
false	true	true	φ_2	[2,1]
false	true	false	φ_5	[1,2]
false	false	true	φ_3	[1,2]
false	false	false	φ_6	[0,3]

b	$\phi(b)$
[3,0]	$\langle\varphi_1\rangle$
[2,1]	$\langle\varphi_4,\varphi_2,\varphi_2\rangle$
[1,2]	$\langle\varphi_5,\varphi_5,\varphi_3\rangle$
[0,3]	$\langle\varphi_6\rangle$

4. Detection of Commutative Factors (DECOR) Algorithm

Properties of buckets:

- For any subset $\boldsymbol{S}$ of commutative arguments:

$$|\boldsymbol{S}| \leq \min_{b \in \{b \mid b \in \mathscr{B}(\phi) \wedge |\phi(b)|>1\}} \max_{\varphi \in \phi(b)} \operatorname{count}(\phi(b),\varphi)$$

  (in each bucket b with $|\phi(b)| > 1$, there is a potential occurring at least $|\boldsymbol{S}|$ times)
- For each group of identical potentials in a bucket, the intersection of their corresponding assignments yields candidates for commutative arguments—e.g., for φ_2's in $[2,1]$:

$$(\mathrm{true},\mathrm{false},\mathrm{true}) \cap (\mathrm{false},\mathrm{true},\mathrm{true}) = (\varnothing,\varnothing,\mathrm{true})$$

A	B	R	$\phi(A,B,R)$	b
true	true	true	φ_1	[3,0]
true	true	false	φ_4	[2,1]
true	false	true	φ_2	[2,1]
true	false	false	φ_5	[1,2]
false	true	true	φ_2	[2,1]
false	true	false	φ_5	[1,2]
false	false	true	φ_3	[1,2]
false	false	false	φ_6	[0,3]

b	$\phi(b)$
[3,0]	$\langle\varphi_1\rangle$
[2,1]	$\langle\varphi_4,\varphi_2,\varphi_2\rangle$
[1,2]	$\langle\varphi_5,\varphi_5,\varphi_3\rangle$
[0,3]	$\langle\varphi_6\rangle$

DECOR iterates over buckets and computes candidates for commutative arguments:

- Initial candidates: $\{\{A,B,R\}\}$
- Candidates for bucket $[3,0]$: Skipped because $|\phi(b)| < 2$
- Candidates for bucket $[2,1]$: $\{\{A,B\}\}$

$$(\mathrm{true},\mathrm{false},\mathrm{true}) \cap (\mathrm{false},\mathrm{true},\mathrm{true}) = (\varnothing,\varnothing,\mathrm{true})$$

- Candidates for bucket $[1,2]$: $\{\{A,B\}\}$

$$(\mathrm{true},\mathrm{false},\mathrm{false}) \cap (\mathrm{false},\mathrm{true},\mathrm{false}) = (\varnothing,\varnothing,\mathrm{false})$$

- Candidates for bucket $[0,3]$: Skipped because $|\phi(b)| < 2$
- Result (intersection over all buckets): $\{\{A,B\}\}$

5. Theoretical Guarantees

Given a factor $\phi(R_1,\ldots,R_n)$:

- DECOR avoids »naive« iteration over all 2^n subsets of arguments
- Time complexity is upper-bounded depending on the number of groups of identical potentials in the buckets
- Expected worst-case time complexity of DECOR is in

$$O\left(|\mathscr{B}(\phi)| \cdot k \cdot \binom{n}{n/2} + r^n \cdot n\right),$$

  where k is the maximum number of potentials within a bucket entailed by ϕ and r is the number of range values of $R_1,\ldots,R_n$

6. Experiments

- Comparison of run times of DECOR and the »naive« approach
- Average run times over factors with $k \in \{0, 2, \lfloor\frac{n}{2}\rfloor, n-1, n\}$ commutative arguments
- Timeout after five minutes per instance