UNIVERSITY OF LÜBECK
INSTITUTE OF INFORMATION SYSTEMS

Efficient Detection of Exchangeable Factors in Factor Graphs

FLAIRS 2024 – Miramar Beach, Florida, USA

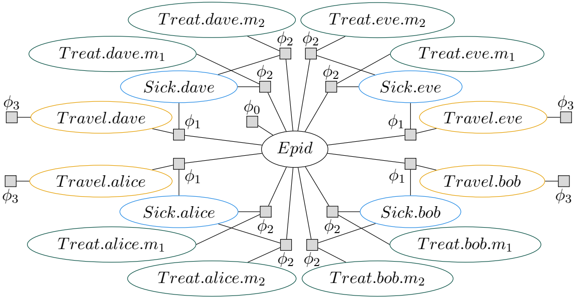
Malte Luttermann[1], Johann Machemer[2], and Marcel Gehrke[2]

[1]German Research Center for Artificial Intelligence (DFKI), Lübeck, Germany
[2]Institute of Information Systems, University of Lübeck, Lübeck, Germany

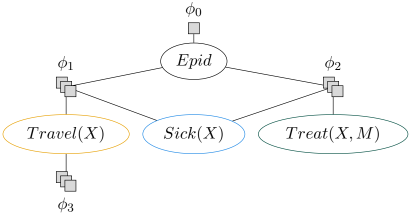
May 19, 2024

Motivation

- Lifting exploits symmetries in factor graphs to speed up probabilistic inference
- Lifting uses a representative of indistinguishable individuals for computations

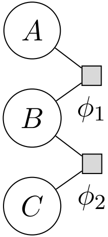


→



Problem Setup I

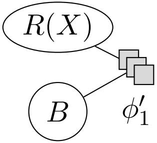
- Goal: Efficiently detect exchangeable factors
- Exchangeable factors encode an equivalent probability distribution



A	B	$\phi_1(A, B)$
true	true	φ_1
true	false	φ_2
false	true	φ_3
false	false	φ_4

C	B	$\phi_2(C, B)$
true	true	φ_1
true	false	φ_2
false	true	φ_3
false	false	φ_4

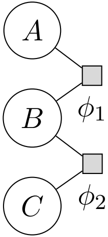
$\rightarrow$



$R(X)$	B	$\phi'_1(R(X), B)$
true	true	φ_1
true	false	φ_2
false	true	φ_3
false	false	φ_4

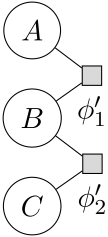
Problem Setup II

- Goal: Efficiently detect exchangeable factors
- Exchangeable factors encode an equivalent probability distribution



A	B	$\phi_1(A, B)$
true	true	φ_1
true	false	φ_2
false	true	φ_3
false	false	φ_4

C	B	$\phi_2(C, B)$
true	true	φ_1
true	false	φ_2
false	true	φ_3
false	false	φ_4



A	B	$\phi'_1(A, B)$
true	true	φ_1
true	false	φ_2
false	true	φ_3
false	false	φ_4

B	C	$\phi'_2(B, C)$
true	true	φ_1
true	false	φ_3
false	true	φ_2
false	false	φ_4

Previous Work and Our Contributions

Previous work (Luttermann et al., 2024):

- Advanced Colour Passing algorithm to construct a lifted representation
- *Buckets* to prune the search space, then iterating over all argument permutations
- Number of iterations in the worst-case for a factor with n arguments: $O(n!)$

Malte Luttermann et al. (2024). »Colour Passing Revisited: Lifted Model Construction with Commutative Factors«. *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-24)*. AAAI Press, pp. 20500–20507.

Previous Work and Our Contributions

Previous work (Luttermann et al., 2024):

- Advanced Colour Passing algorithm to construct a lifted representation
- *Buckets* to prune the search space, then iterating over all argument permutations
- Number of iterations in the worst-case for a factor with n arguments: $O(n!)$

Our contributions:

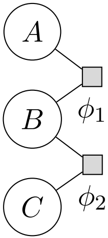
- Theoretical guarantees: *Buckets* to avoid iterating over permutations if possible
- Practical algorithm: Detection of Exchangeable Factors (DEFT) algorithm
- Number of iterations in the worst-case for a factor with n arguments: $O(d)$ with $d \ll n!$ in many practical settings

---

Malte Luttermann et al. (2024). »Colour Passing Revisited: Lifted Model Construction with Commutative Factors«. *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-24)*. AAAI Press, pp. 20500–20507.

Buckets

- Buckets count the occurrences of specific range values in an assignment
- Each potential belongs to *exactly one* bucket
- Each bucket contains *at least one* potential



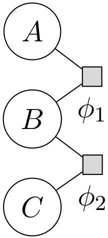
A	B	$\phi_1(A, B)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_2	$[1, 1]$
false	true	φ_3	$[1, 1]$
false	false	φ_4	$[0, 2]$

B	C	$\phi_2(B, C)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_3	$[1, 1]$
false	true	φ_2	$[1, 1]$
false	false	φ_4	$[0, 2]$

b	$\phi_1(b)$	$\phi_2(b)$
$[2, 0]$	$\{\varphi_1\}$	$\{\varphi_1\}$
$[1, 1]$	$\{\varphi_2, \varphi_3\}$	$\{\varphi_3, \varphi_2\}$
$[0, 2]$	$\{\varphi_4\}$	$\{\varphi_4\}$

Properties of Buckets I

- If at least one bucket is mapped to different multisets of values by ϕ_1 and ϕ_2, then ϕ_1 and ϕ_2 are not exchangeable (Luttermann et al., 2024)



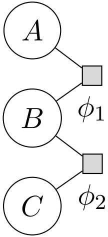
A	B	$\phi_1(A, B)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_2	$[1, 1]$
false	true	φ_3	$[1, 1]$
false	false	φ_4	$[0, 2]$

B	C	$\phi_2(B, C)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_3	$[1, 1]$
false	true	φ_2	$[1, 1]$
false	false	φ_4	$[0, 2]$

b	$\phi_1(b)$	$\phi_2(b)$
$[2, 0]$	$\{\varphi_1\}$	$\{\varphi_1\}$
$[1, 1]$	$\{\varphi_2, \varphi_3\}$	$\{\varphi_3, \varphi_2\}$
$[0, 2]$	$\{\varphi_4\}$	$\{\varphi_4\}$

Properties of Buckets II

- Two factors are exchangeable iff there exists a permutation of their arguments such that each bucket is mapped to the same ordered multiset of values



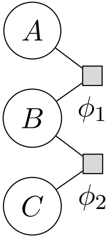
A	B	$\phi_1(A, B)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_2	$[1, 1]$
false	true	φ_3	$[1, 1]$
false	false	φ_4	$[0, 2]$

B	C	$\phi_2(B, C)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_3	$[1, 1]$
false	true	φ_2	$[1, 1]$
false	false	φ_4	$[0, 2]$

b	$\phi_1^{\succ}(b)$	$\phi_2^{\succ}(b)$
$[2, 0]$	$\langle \varphi_1 \rangle$	$\langle \varphi_1 \rangle$
$[1, 1]$	$\langle \varphi_2, \varphi_3 \rangle$	$\langle \varphi_3, \varphi_2 \rangle$
$[0, 2]$	$\langle \varphi_4 \rangle$	$\langle \varphi_4 \rangle$

The DEFT Algorithm

- Iterate over buckets and search for possible rearrangements to obtain identically ordered multisets within each bucket



A	B	$\phi_1(A, B)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_2	$[1, 1]$
false	true	φ_3	$[1, 1]$
false	false	φ_4	$[0, 2]$

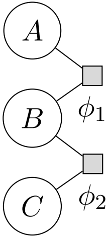
B	C	$\phi_2(B, C)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_3	$[1, 1]$
false	true	φ_2	$[1, 1]$
false	false	φ_4	$[0, 2]$

b	$\phi_1^{\succ}(b)$	$\phi_2^{\succ}(b)$
$[2, 0]$	$\langle \varphi_1 \rangle$	$\langle \varphi_1 \rangle$
$[1, 1]$	$\langle \varphi_2, \varphi_3 \rangle$	$\langle \varphi_3, \varphi_2 \rangle$
$[0, 2]$	$\langle \varphi_4 \rangle$	$\langle \varphi_4 \rangle$

Possible rearrangements:
$[2, 0]$: $B \to \{1, 2\}$, $C \to \{1, 2\}$

The DEFT Algorithm

- Iterate over buckets and search for possible rearrangements to obtain identically ordered multisets within each bucket



A	B	$\phi_1(A,B)$	b
true	true	φ_1	$[2,0]$
true	false	φ_2	$[1,1]$
false	true	φ_3	$[1,1]$
false	false	φ_4	$[0,2]$

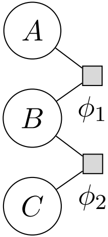
B	C	$\phi_2(B,C)$	b
true	true	φ_1	$[2,0]$
true	false	φ_3	$[1,1]$
false	true	φ_2	$[1,1]$
false	false	φ_4	$[0,2]$

b	$\phi_1^{\succ}(b)$	$\phi_2^{\succ}(b)$
$[2,0]$	$\langle \varphi_1 \rangle$	$\langle \varphi_1 \rangle$
$[1,1]$	$\langle \varphi_2, \varphi_3 \rangle$	$\langle \varphi_3, \varphi_2 \rangle$
$[0,2]$	$\langle \varphi_4 \rangle$	$\langle \varphi_4 \rangle$

Possible rearrangements:
$[2,0]$: $B \to \{1,2\}$, $C \to \{1,2\}$
$[1,1]$: $B \to \{2\}$, $C \to \{1\}$

The DEFT Algorithm

- Iterate over buckets and search for possible rearrangements to obtain identically ordered multisets within each bucket



A	B	$\phi_1(A, B)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_2	$[1, 1]$
false	true	φ_3	$[1, 1]$
false	false	φ_4	$[0, 2]$

B	C	$\phi_2(B, C)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_3	$[1, 1]$
false	true	φ_2	$[1, 1]$
false	false	φ_4	$[0, 2]$

b	$\phi_1^{\succ}(b)$	$\phi_2^{\succ}(b)$
$[2, 0]$	$\langle \varphi_1 \rangle$	$\langle \varphi_1 \rangle$
$[1, 1]$	$\langle \varphi_2, \varphi_3 \rangle$	$\langle \varphi_3, \varphi_2 \rangle$
$[0, 2]$	$\langle \varphi_4 \rangle$	$\langle \varphi_4 \rangle$

Possible rearrangements:

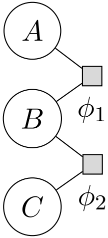
$[2, 0]$: $B \to \{1, 2\}$, $C \to \{1, 2\}$

$[1, 1]$: $B \to \{2\}$, $C \to \{1\}$

$[0, 2]$: $B \to \{1, 2\}$, $C \to \{1, 2\}$

The DEFT Algorithm

- Iterate over buckets and search for possible rearrangements to obtain identically ordered multisets within each bucket



A	B	$\phi_1(A, B)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_2	$[1, 1]$
false	true	φ_3	$[1, 1]$
false	false	φ_4	$[0, 2]$

B	C	$\phi_2(B, C)$	b
true	true	φ_1	$[2, 0]$
true	false	φ_3	$[1, 1]$
false	true	φ_2	$[1, 1]$
false	false	φ_4	$[0, 2]$

b	$\phi_1^{\succ}(b)$	$\phi_2^{\succ}(b)$
$[2, 0]$	$\langle \varphi_1 \rangle$	$\langle \varphi_1 \rangle$
$[1, 1]$	$\langle \varphi_2, \varphi_3 \rangle$	$\langle \varphi_3, \varphi_2 \rangle$
$[0, 2]$	$\langle \varphi_4 \rangle$	$\langle \varphi_4 \rangle$

Possible rearrangements:
$[2, 0]$: $B \to \{1, 2\}$, $C \to \{1, 2\}$
$[1, 1]$: $B \to \{2\}$, $C \to \{1\}$
$[0, 2]$: $B \to \{1, 2\}$, $C \to \{1, 2\}$
Intersection: $B \to \{2\}$, $C \to \{1\}$

Theoretical Guarantees of the DEFT Algorithm

- Duplicate values in buckets increase complexity
- Degree of freedom of a bucket b:

$$\mathcal{F}(b) = \prod_{\varphi \in \text{unique}(\phi^{\succ}(b))} \text{count}(\phi^{\succ}(b), \varphi)!$$

- Degree of freedom of a factor ϕ:

$$\mathcal{F}(\phi) = \min_{b \in \{b \mid b \in \mathcal{B}(\phi) \land |\phi^{\succ}(b)| > 1\}} \mathcal{F}(b)$$

b	$\phi_1^{\prime\succ}(b)$	$\phi_2^{\prime\succ}(b)$
$[3,0]$	$\langle \varphi_1 \rangle$	$\langle \varphi_1 \rangle$
$[2,1]$	$\langle \varphi_2, \varphi_3, \varphi_5 \rangle$	$\langle \varphi_3, \varphi_5, \varphi_2 \rangle$
$[1,2]$	$\langle \varphi_4, \varphi_6, \varphi_6 \rangle$	$\langle \varphi_6, \varphi_4, \varphi_6 \rangle$
$[0,3]$	$\langle \varphi_7 \rangle$	$\langle \varphi_7 \rangle$

Theoretical Guarantees of the DEFT Algorithm

- Duplicate values in buckets increase complexity
- Degree of freedom of a bucket b:

$$\mathcal{F}(b) = \prod_{\varphi \in \text{unique}(\phi^{\succ}(b))} \text{count}(\phi^{\succ}(b), \varphi)!$$

b	$\phi_1^{\prime\succ}(b)$	$\phi_2^{\prime\succ}(b)$
$[3,0]$	$\langle \varphi_1 \rangle$	$\langle \varphi_1 \rangle$
$[2,1]$	$\langle \varphi_2, \varphi_3, \varphi_5 \rangle$	$\langle \varphi_3, \varphi_5, \varphi_2 \rangle$
$[1,2]$	$\langle \varphi_4, \varphi_6, \varphi_6 \rangle$	$\langle \varphi_6, \varphi_4, \varphi_6 \rangle$
$[0,3]$	$\langle \varphi_7 \rangle$	$\langle \varphi_7 \rangle$

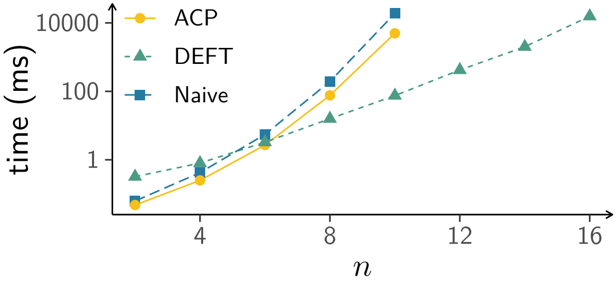
- Degree of freedom of a factor ϕ:

$$\mathcal{F}(\phi) = \min_{b \in \{b \mid b \in \mathcal{B}(\phi) \wedge |\phi^{\succ}(b)| > 1\}} \mathcal{F}(b)$$

- The number of table comparisons needed to check whether ϕ_1 and ϕ_2 are exchangeable is in $O(d)$, where $d = \min\{\mathcal{F}(\phi_1), \mathcal{F}(\phi_2)\}$ (i.e., d is factorial)
- In many practical settings it holds that $d \ll n!$
- The degree of freedom is upper-bounded by $n!$

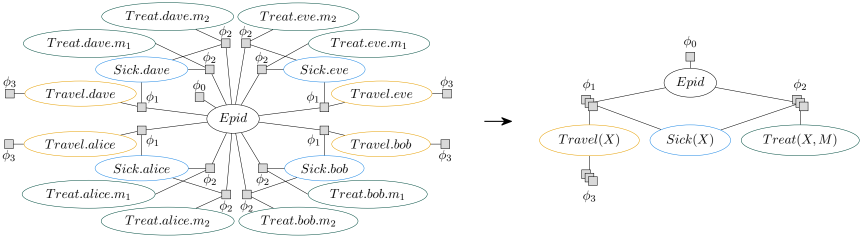
Experiments

- Comparison of run times of DEFT, ACP (previous work), and a »naive« approach
- Average over exchangeable and non-exchangeable factors with a proportion of identical potentials in $\{0.0, 0.1, 0.2, 0.5, 0.8, 0.9, 1.0\}$
- Timeout after 30 minutes per instance



Summary

- Problem of detecting exchangeable factors efficiently solved
- Upper bound on the number of table comparisons depending on the number of duplicate potential values within the buckets of the factors
- The DEFT algorithm exploits this upper bound effectively in practice



References

Malte Luttermann, Tanya Braun, Ralf Möller, and Marcel Gehrke (2024). »Colour Passing Revisited: Lifted Model Construction with Commutative Factors«. *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-24)*. AAAI Press, pp. 20500–20507.