dfki ai
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

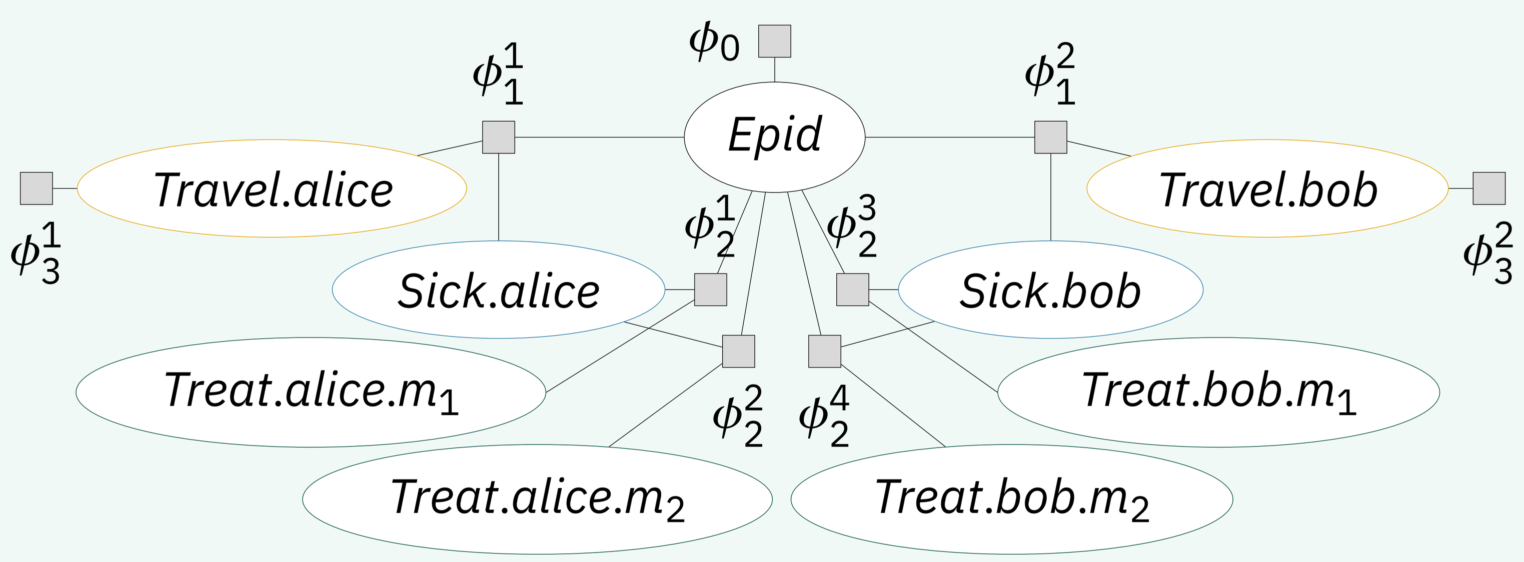
Lifted Model Construction without Normalisation: A Vectorised Approach to Exploit Symmetries in Factor Graphs

Malte Luttermann, Ralf Möller, and Marcel Gehrke

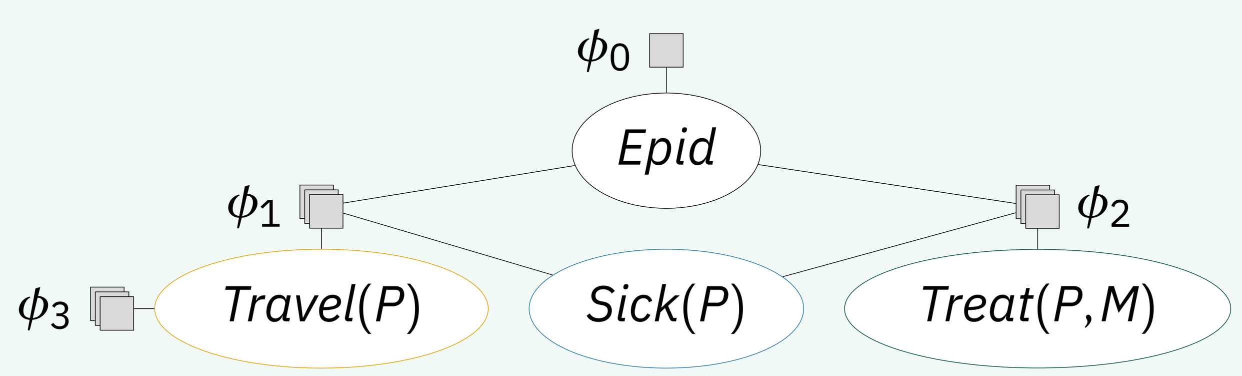
1. Motivation and Problem Setup

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

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



- Parametric factor graphs introduce logical variables to represent groups of random variables



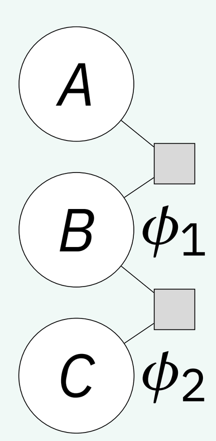
Problem Setup

Input: A factor graph G
Output: A parametric factor graph equivalent to G
Constraint: Handle different scaling of factors

- Factors with different scaling can occur when learning a factor graph from observed data
  - Occurrences of range values are counted
  - Normalisation is not always possible / desirable (e.g., to avoid floating point arithmetics)

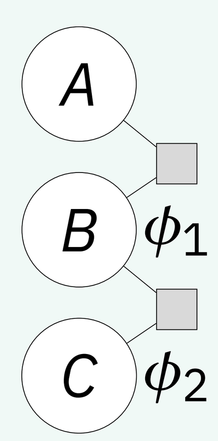
2. Exchangeable Factors

- Exchangeable factors are identical up to scalar $\alpha \in \mathbb{R}^+$
- Exchangeable factors can be grouped



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



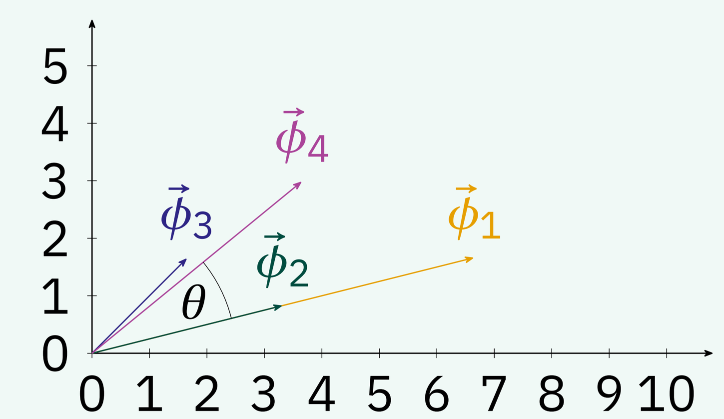
A	B	$\phi_1(A,B)$
true	true	φ_1
true	false	φ_2
false	true	φ_3
false	false	φ_4

C	B	$\phi_2(C,B)$
true	true	$\alpha \cdot \varphi_1$
true	false	$\alpha \cdot \varphi_2$
false	true	$\alpha \cdot \varphi_3$
false	false	$\alpha \cdot \varphi_4$

- Theorem: Scaling any factor by $\alpha \in \mathbb{R}^+$ leaves the semantics of the factor graph unchanged

3. Vector Representation of Factors

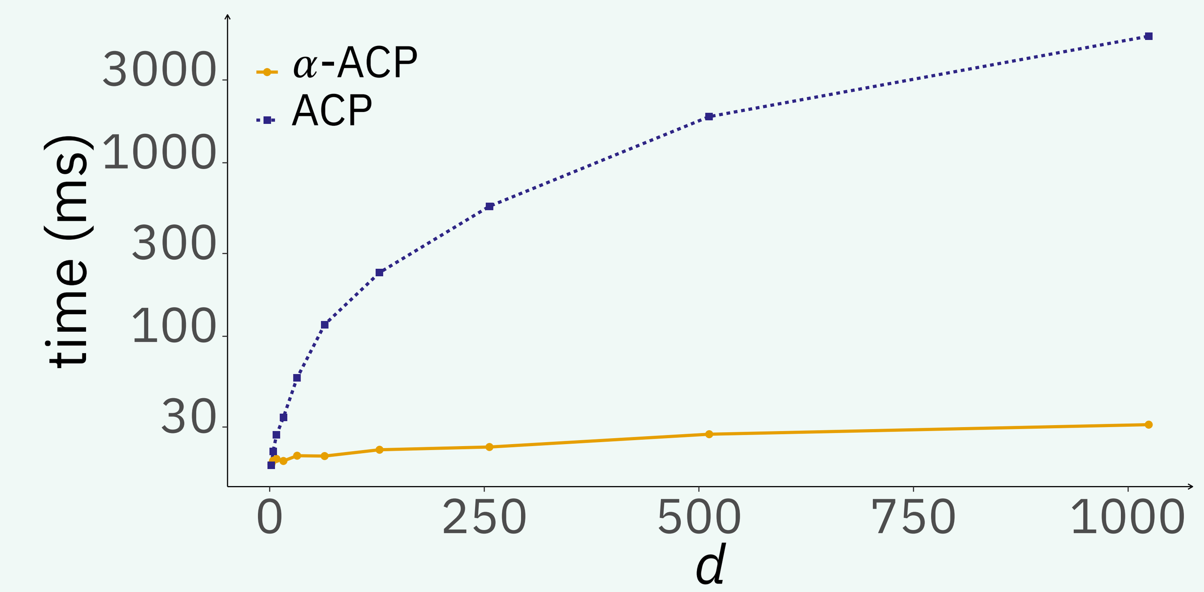
- Idea: Represent table of potentials as a vector
- Theorem: Exchangeable factors have collinear vectors
- E.g., two dimensions:
- ϕ_1, ϕ_2 are exchangeable



- Collinearity can be checked without any division operators to avoid floating point arithmetics
- Approach can also be extended:
  (1) To allow for permuted arguments
  (2) To allow for approximate exchangeability

4. Experiments

- Comparison of run times for probabilistic inference
  - ACP is the previous state-of-the-art algorithm
  - α-ACP is our generalisation of ACP
  - d controls the size of the input factor graph
- Average time over factor graphs where a proportion of $p \in \{0.01, 0.05, 0.1, 0.15\}$ of factors is scaled
- ACP is not able to detect exchangeable factors with different scales, whereas α-ACP is



- β is average number of queries after which offline overhead of α-ACP compared to ACP amortises
- Negative values for β are omitted as there is no overhead in this case (hence the missing boxes)
- After at most 10 queries, the overhead amortises