

Qualitative Chain Graphs and their Use in Medicine

Martijn Lappenschaar, Arjen Hommersom, Peter J.F. Lucas

Institute for Computing and Information Sciences, Radboud University Nijmegen, The Netherlands
 {mlappens1, arjenh, peterl}@cs.ru.nl

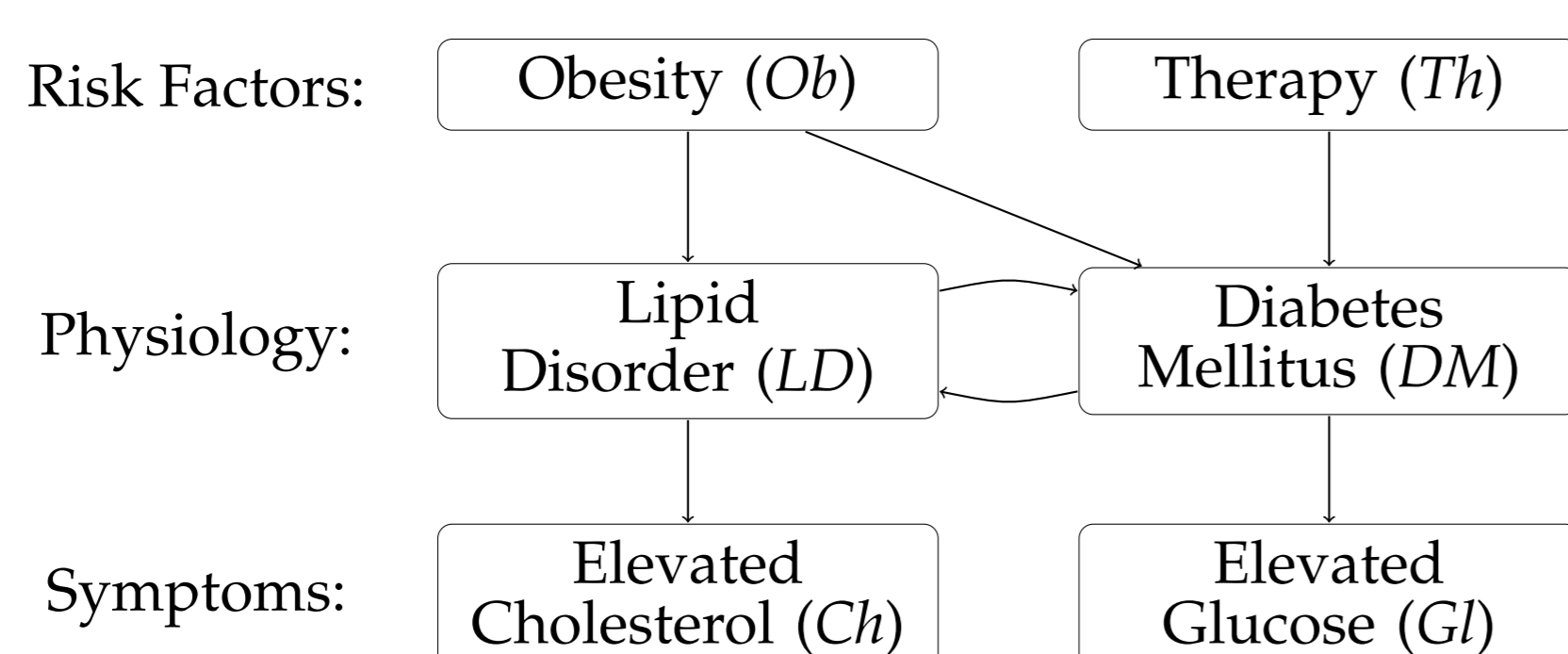
Outline

For modelling diseases in medicine, chain graphs are more attractive than directed graphs, i.e., Bayesian networks, as they support representing interactions between diseases that have no natural direction; in particular, they can be used to model **equilibria**. However, it is difficult to interpret chain graph parameters, making them more **blackbox** models compared to Bayesian networks. In this paper, we developed a **qualitative abstraction** of chain graph parameters, making these models more suitable as whitebox models. Further, by looking upon these abstractions as **constraints on the probability distribution**, we derive distributions over marginals. We show that this can be useful for **decision making** in complex medical domains.

Motivation: feedback systems in medicine

Many physiological processes within the human body can be seen as **causal feedback systems**, in which some kind of equilibrium setpoint is maintained. In non-healthy people the **equilibrium setpoint** typically differs from the healthy people, but therapeutic interventions can reset the equilibrium setpoint to a state that is closer to the healthy people. The **disturbance of the equilibrium** of one physiological process, might also alter the equilibrium setpoints of **other regulation systems**, which might in turn induce new pathophysiology and decrease the patient's prognosis even further.

Consider, for example, the following schematic representation of an interaction between diabetes mellitus and a lipid disorder, suggesting a feedback mechanism between their pathophysiology.



Chain graph representation

This medical example in its equilibrium can be modelled as a chain graph.

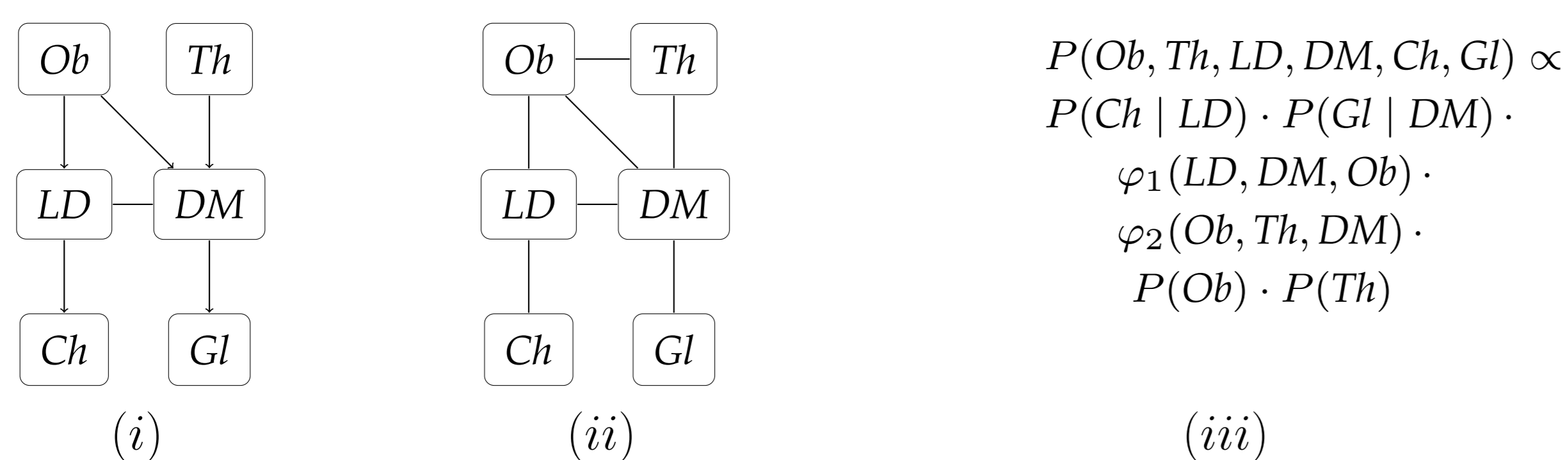


Figure 1: Chain graph representation (i), closure graph of chain components (ii), and factorisation (iii) of the example.

Qualitative Chain Graphs

We developed a qualitative representation consisting of:

- Qualitative influences $S(A, B)$: the effect of a cause A on B
- Qualitative synergies: interaction of two causes on an effect, including:
 - Additive synergy $Y(A_1, A_2, B)$
 - Product synergy $X(A_1, A_2, B)$
- Probabilistic relationships have signs: +, -, 0, or ?

For example:

- A vertex A positively influences a vertex B , written as $S^+(A, B)$, iff $A \in \text{bd}(B)$ and

$$P(b | a, \text{bd}(B) - A) \geq P(b | \bar{a}, \text{bd}(B) - A)$$

where $\text{bd}(B)$ is the *boundary* (parents and neighbours) of B .

- Vertices A_1 and A_2 have a positive additive synergy on a vertex B , written as $Y^+(\{A_1, A_2\}, B)$, iff $A_1, A_2 \in \text{bd}(B)$, $Z = \text{bd}(B) - A_1A_2$, and

$$P(b | a_1, a_2, Z) - P(b | \bar{a}_1, a_2, Z) \geq P(b | a_1, \bar{a}_2, Z) - P(b | \bar{a}_1, \bar{a}_2, Z)$$

- Vertices A_1 and A_2 have a negative product synergy with regard to the value b on the vertex B , written as $X^-(\{A_1, A_2\}, b)$, iff $A_1, A_2 \in \text{bd}(B)$, $Z = \text{bd}(B) - A_1A_2$, and

$$P(b | a_1, a_2, Z) \cdot P(b | \bar{a}_1, \bar{a}_2, Z) \geq P(b | a_1, \bar{a}_2, Z) \cdot P(b | \bar{a}_1, a_2, Z)$$

Inference

Influences and synergies define *local constraints* on probability distributions, e.g.

$$P(b | a, \text{bd}(B) - A) \geq P(b | \bar{a}, \text{bd}(B) - A)$$

if and only if

$$\prod_{M \in M_{AB}} \varphi_M(a, b) \varphi_M(\bar{a}, \bar{b}) \geq \prod_{M \in M_{AB}} \varphi_M(a, \bar{b}) \varphi_M(\bar{a}, b)$$

where $M_{AB} = \{M \in M_C \mid \{A, B\} \subseteq M\}$.

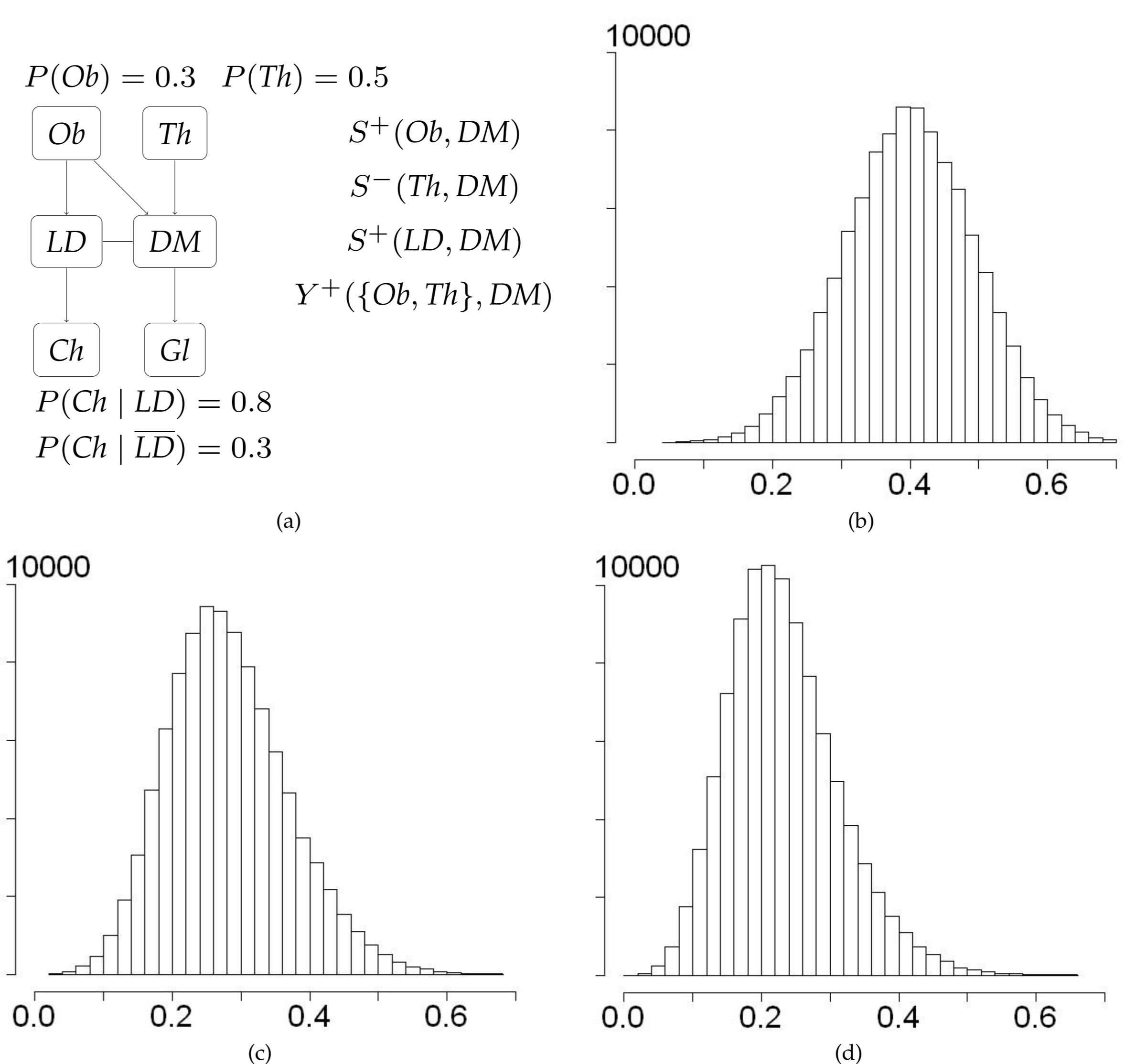
Such properties can be exploited when sampling distributions as follows.

Input: potentials, ϕ_{known} , ϕ_{unknown} , qualitative constraints C

```

for  $\phi_M \in \phi_{\text{unknown}}$  do
   $\phi_M \leftarrow \text{sample a potential for variables } M$ 
   $\phi_{\text{known}} \leftarrow \phi_{\text{known}} \cup \{\phi_M\}$ 
  while !satisfies( $\phi_{\text{known}}, C$ ) do
    resample potentials in failed local constraints
  end while
end for
return distribution based on  $\phi_{\text{known}}$ 
    
```

Experiments



(a) Model with quantitative and qualitative information.

(b) Distribution of $P(Ch)$ (high cholesterol).

(c) An intervention on Th leads to lower probabilities for Ch . The probability $P(Ch | Th) < P(Ch) \approx 0.82$ within the generated samples, showing with a fairly high confidence that diabetic therapy is also beneficial to reduce cholesterol levels.

(d) An additional positive synergy between Ob and Th on DM pushes the distribution even more to lower probabilities with an intervention on Ob . The probability $P(Ch | \bar{Ob}, Th) < P(Ch) \approx 0.91$, suggesting that an additional reduction of weight in combination with diabetic therapy is even more beneficial to reduce cholesterol levels.

Future Work

- Application to real-world multimorbidity data.
- Sampling potentials using more advanced strategies.