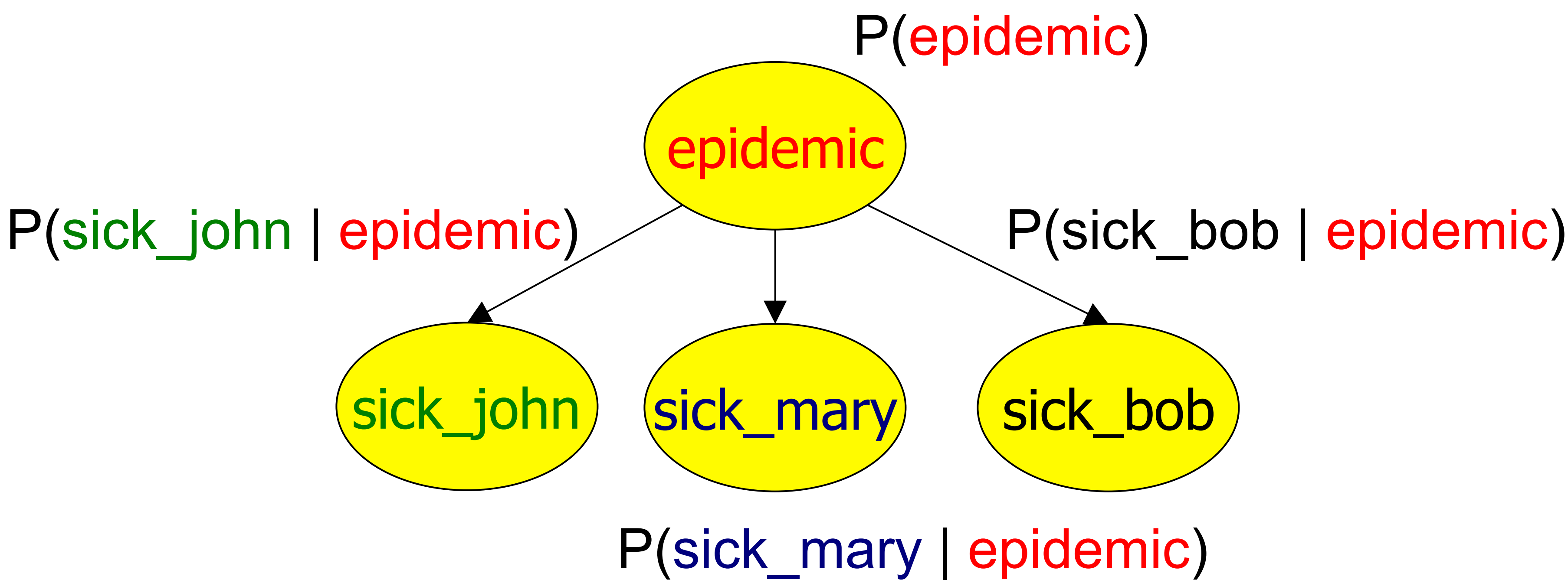
Lifted First-Order Probabilistic Inference

Rodrigo de Salvo Braz
SRI International
based on work with
Eyal Amir and Dan Roth

Outline

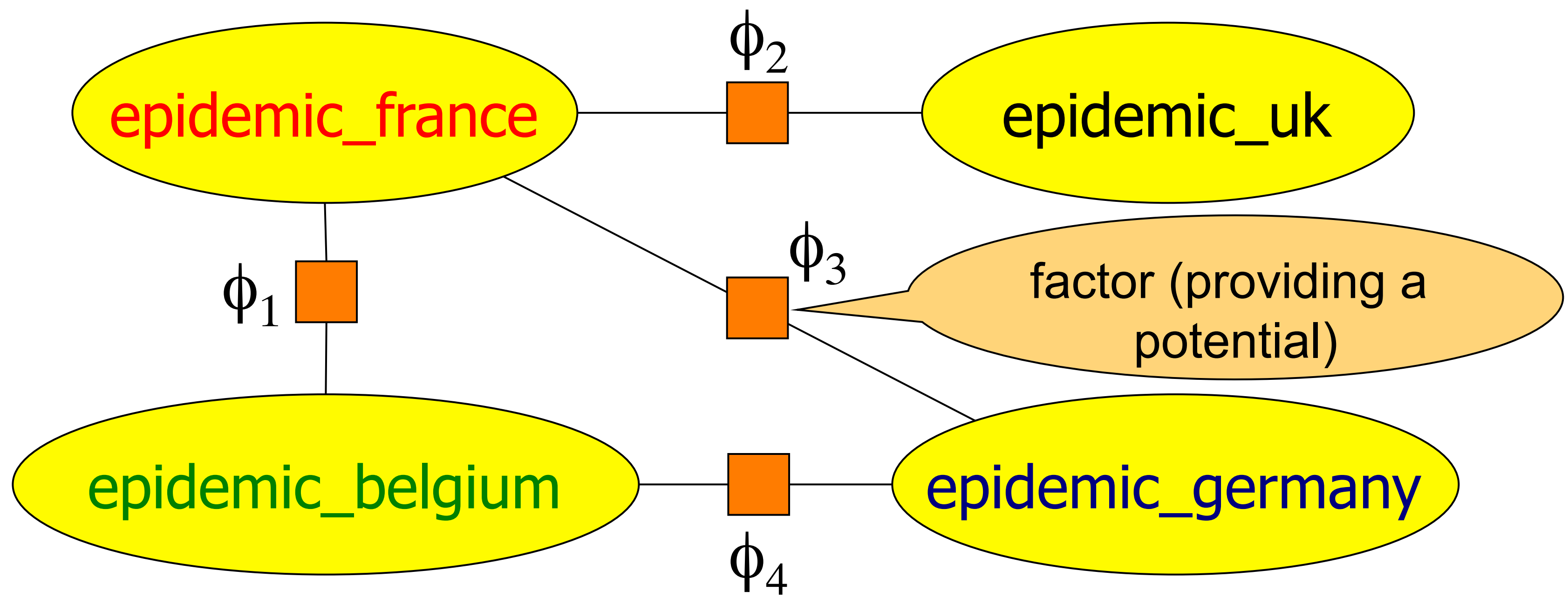
- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion

Bayesian Networks (directed)



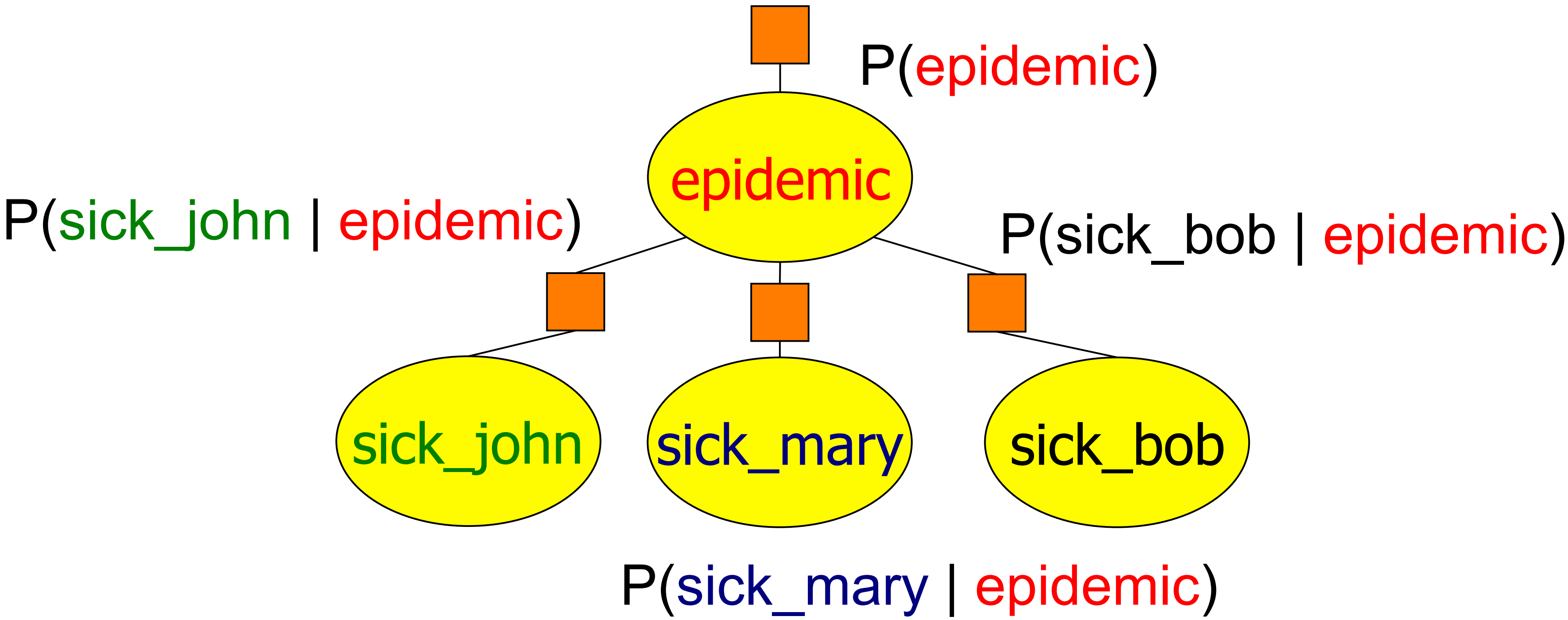
P(sick_john, sick_mary, sick_bob, epidemic)
= P(sick_john | epidemic) * P(sick_mary | epidemic)
* P(sick_bob | epidemic) * P(epidemic)

Factor Networks (undirected)



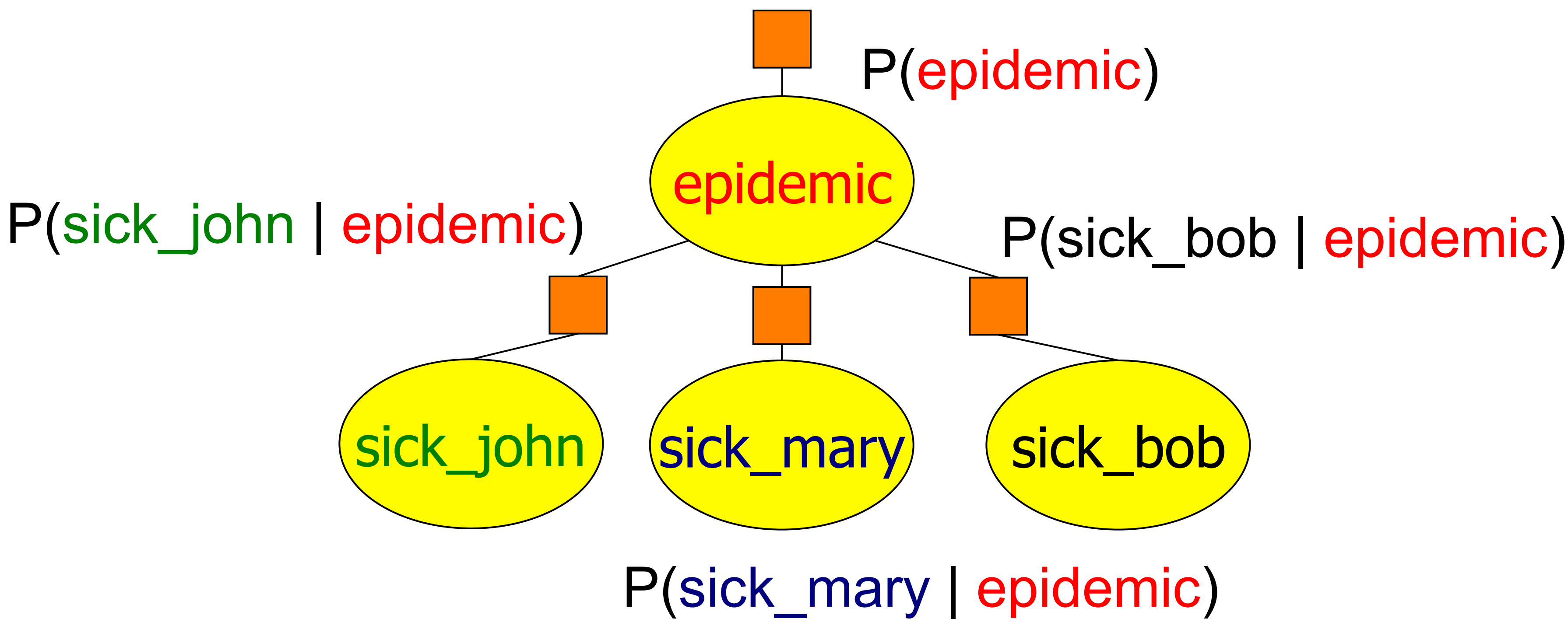
P(epi_france, epi_belgium, epi_uk, epi_germany)
/ ϕ_1(epi_france, epi_belgium) * ϕ_2(epi_france, epi_uk)
* ϕ_3(epi_france, epi_germany) * ϕ_4(epi_belgium, epi_germany)

Bayesian Nets as Factor Networks



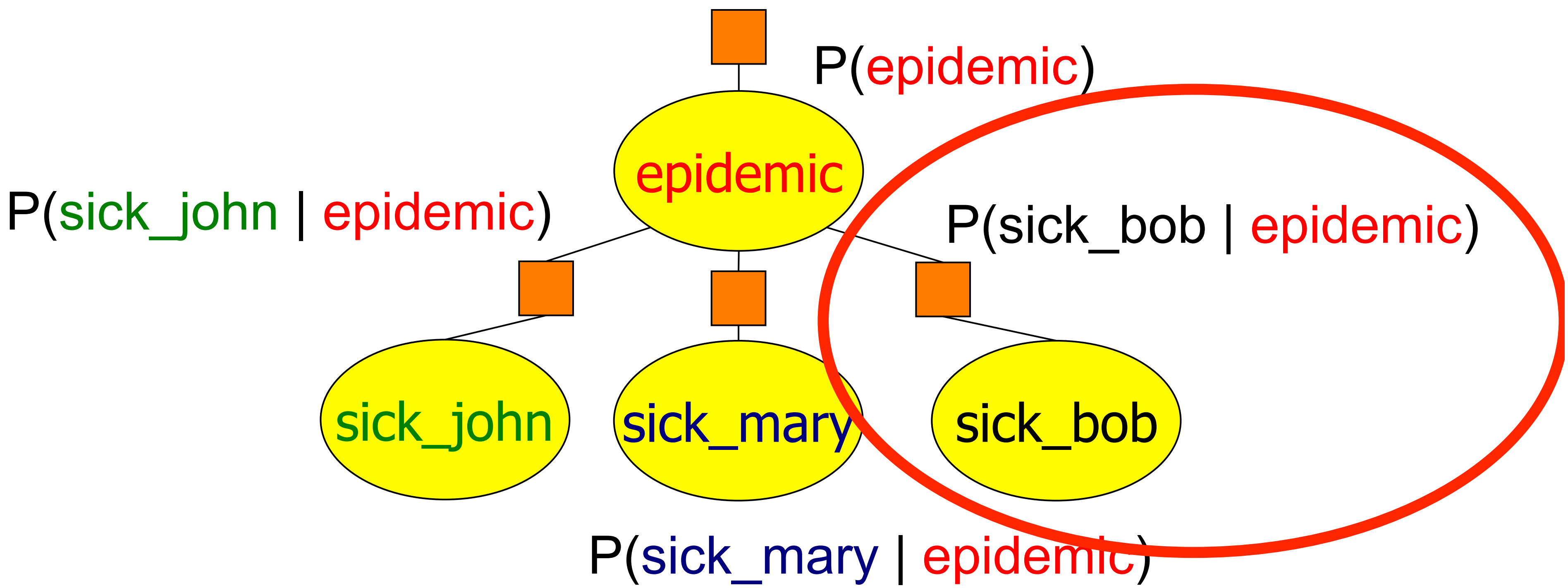
P(sick_john, sick_mary, sick_bob, epidemic)
/ P(sick_john | epidemic) * P(sick_mary | epidemic)
* P(sick_bob | epidemic) * P(epidemic)

Inference: Marginalization



$$P(\text{sick\_john}) \;/\; \sum_{\text{epidemic}} \sum_{\text{sick\_mary}} \sum_{\text{sick\_bob}} P(\text{sick\_john} \mid \text{epidemic}) * P(\text{sick\_mary} \mid \text{epidemic}) * P(\text{sick\_bob} \mid \text{epidemic}) * P(\text{epidemic})$$

Inference: Variable Elimination (VE)

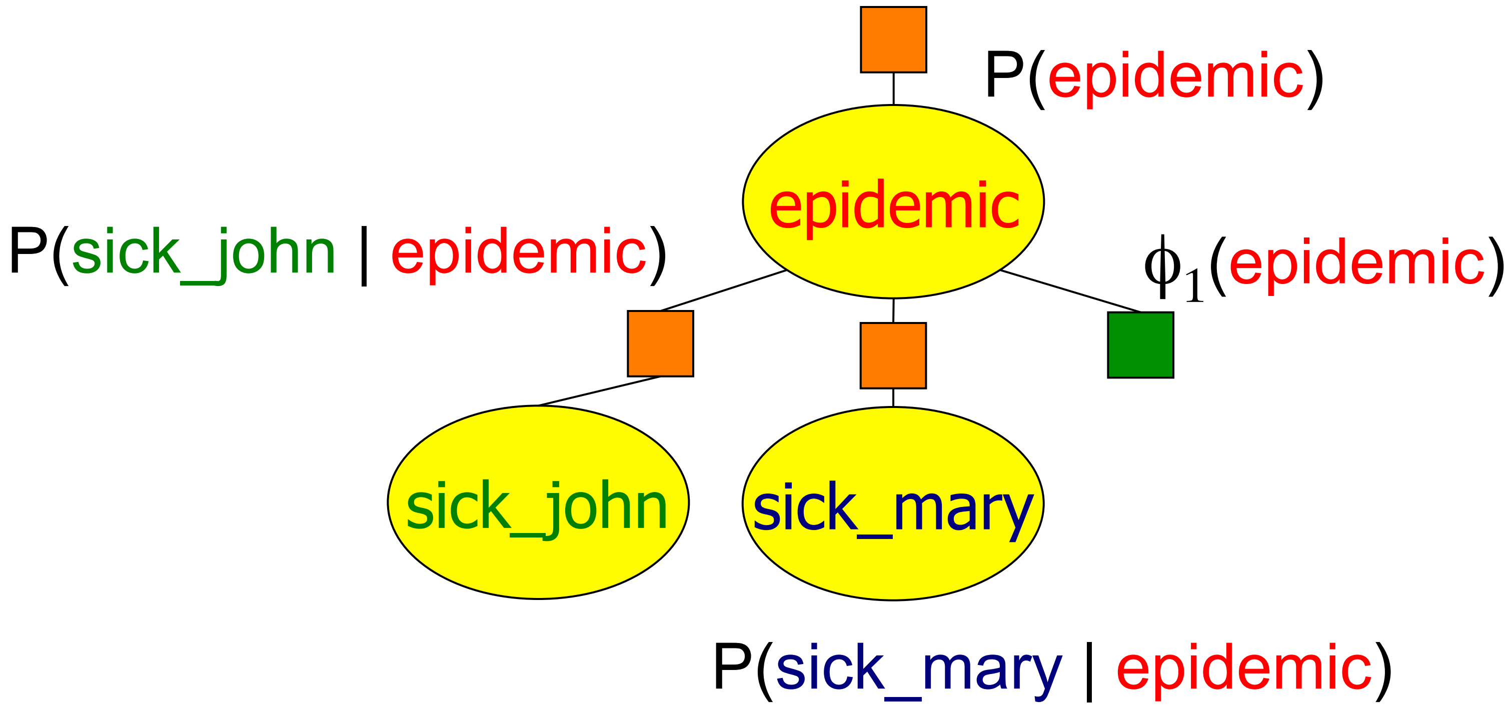


$$P(\text{sick\_john}) \,/\, \sum_{\text{epidemic}} P(\text{sick\_john} \mid \text{epidemic}) * P(\text{epidemic})$$

$$* \sum_{\text{sick\_mary}} P(\text{sick\_mary} \mid \text{epidemic})$$

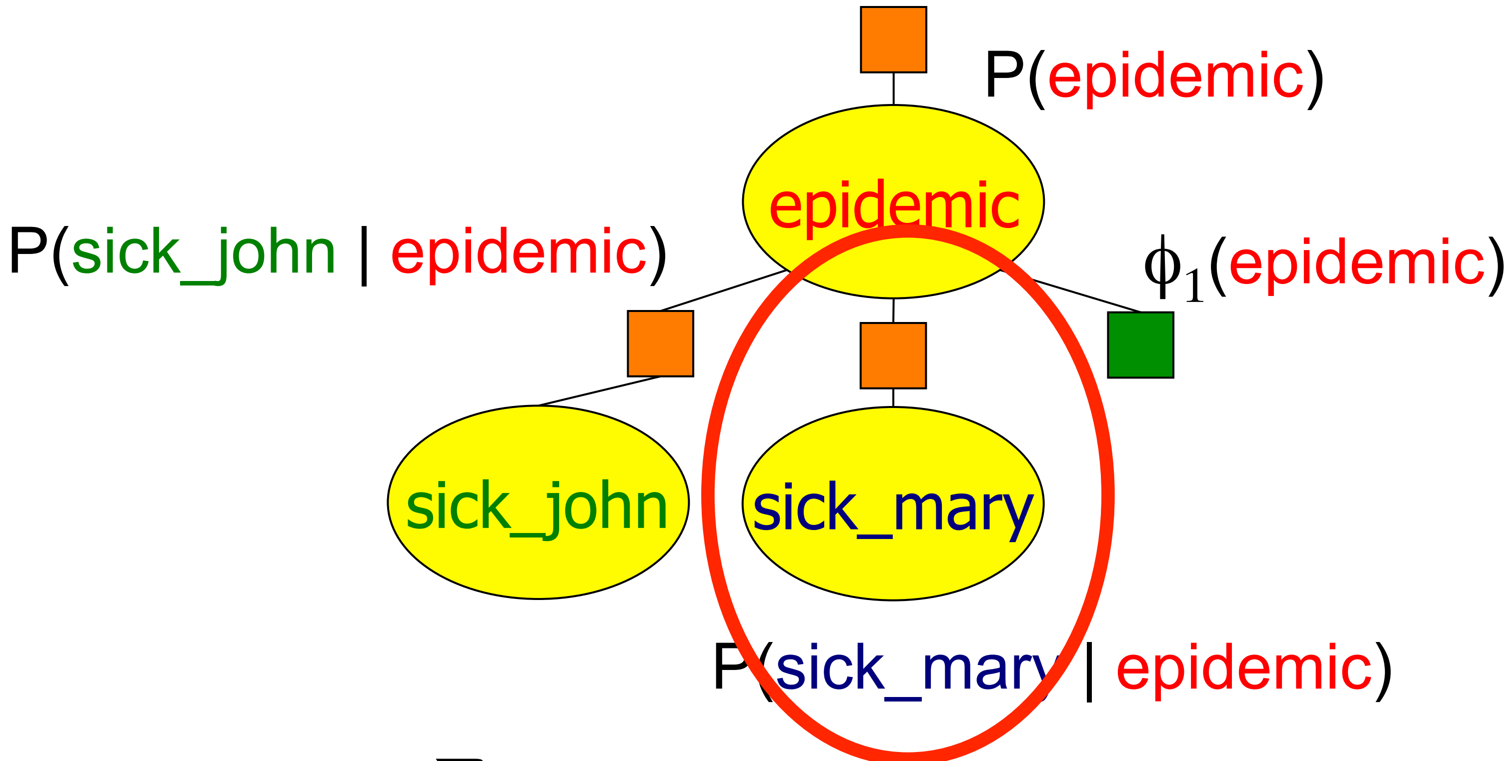
$$* \sum_{\text{sick\_bob}} P(\text{sick\_bob} \mid \text{epidemic})$$

Inference: Variable Elimination (VE)



$$P(\text{sick\_john}) \;/\; \sum_{\text{epidemic}} P(\text{sick\_john} \mid \text{epidemic}) * P(\text{epidemic}) * \sum_{\text{sick\_mary}} P(\text{sick\_mary} \mid \text{epidemic}) * \phi_1(\text{epidemic})$$

Inference: Variable Elimination (VE)

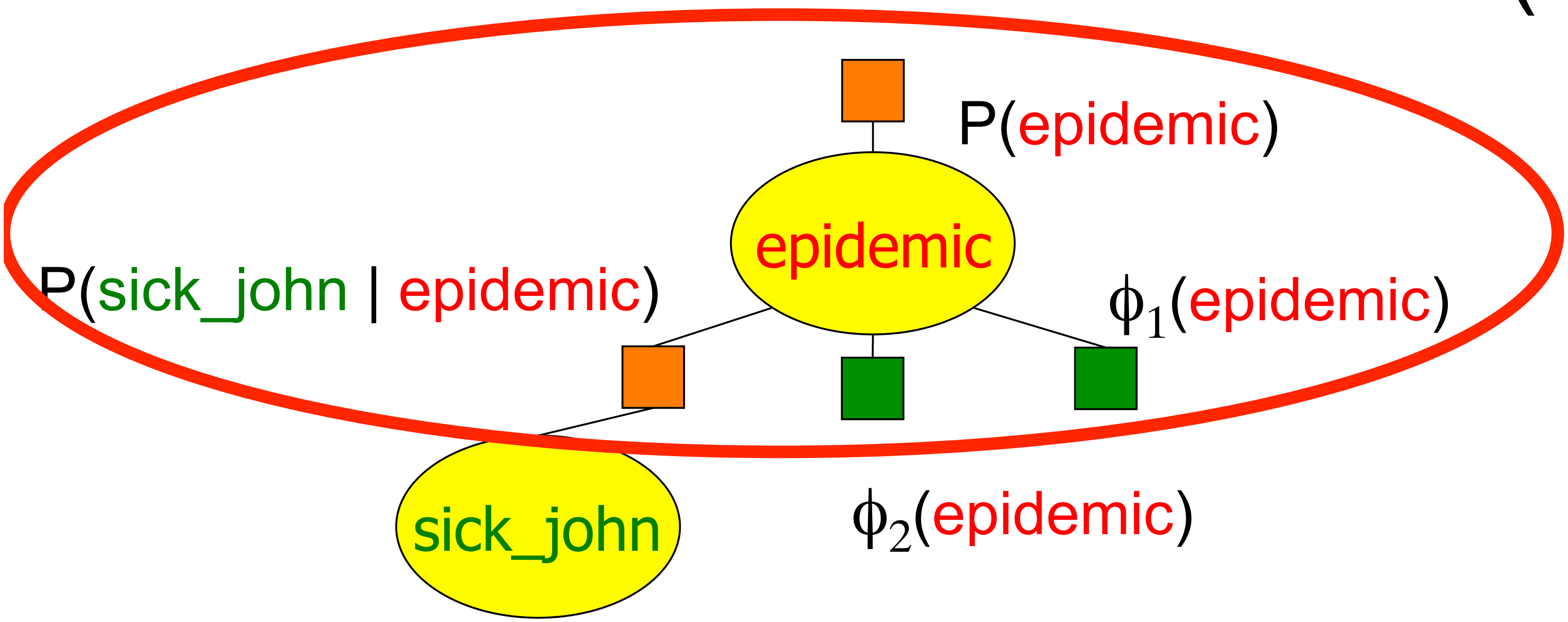


$P(sick\_john) / \sum_{epidemic} P(sick\_john \mid epidemic) * P(epidemic)$

$* \phi_1 (epidemic)$

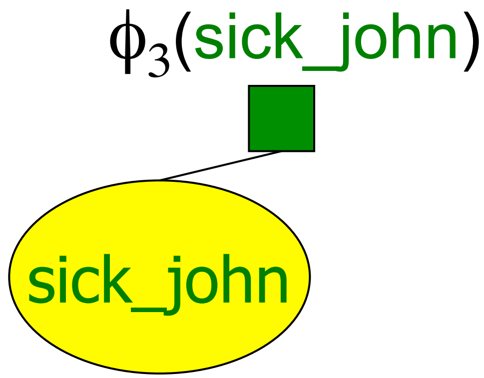
$* \sum_{sick\_mary} P(sick\_mary \mid epidemic)$

Inference: Variable Elimination (VE)



$$P(\text{sick\_john}) \,/\, \sum_{\text{epidemic}} P(\text{sick\_john} \mid \text{epidemic}) * P(\text{epidemic}) * \phi_1(\text{epidemic}) * \phi_2(\text{epidemic})$$

Inference: Variable Elimination (VE)



P(sick_john) / ϕ_3 (sick_john)

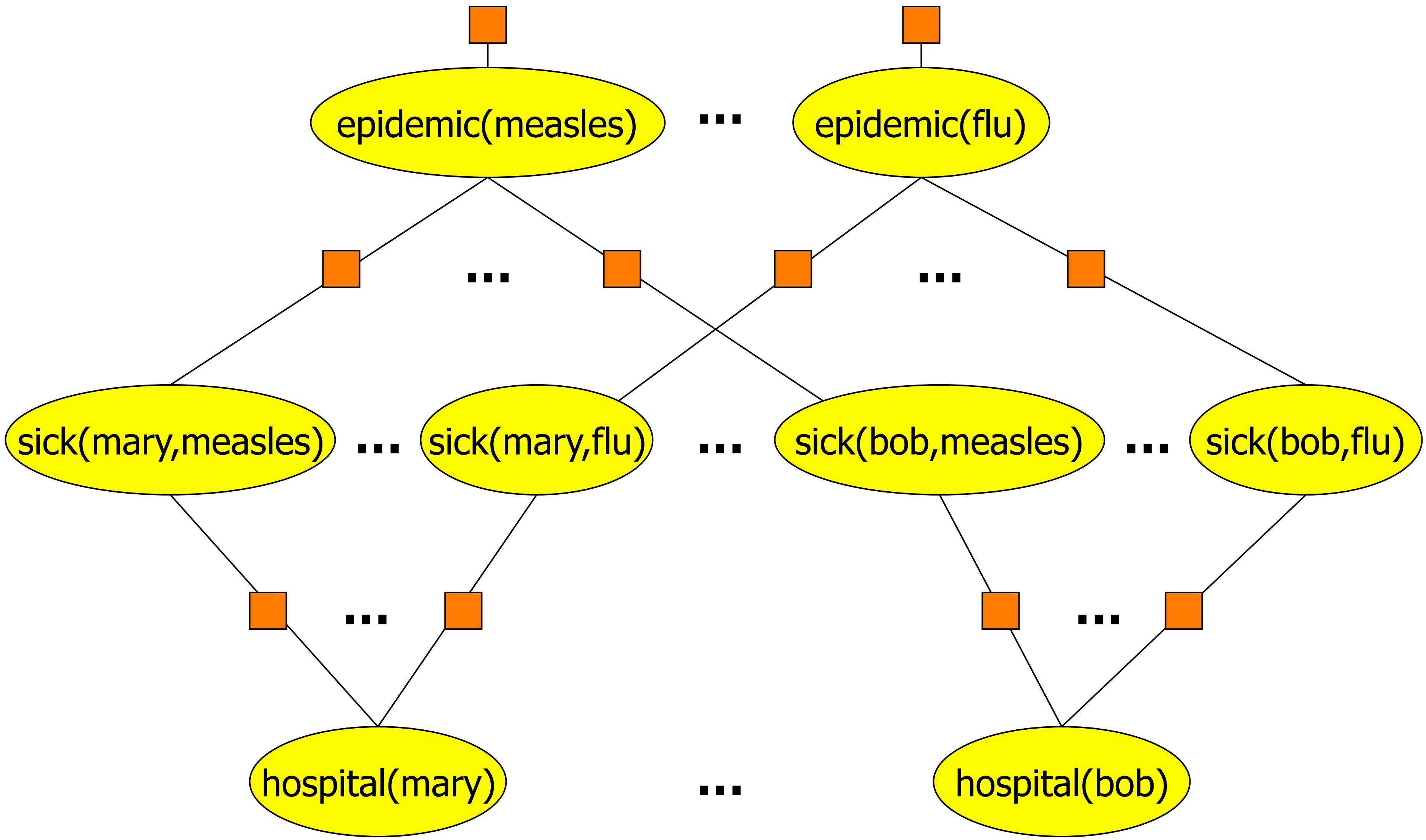
Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion

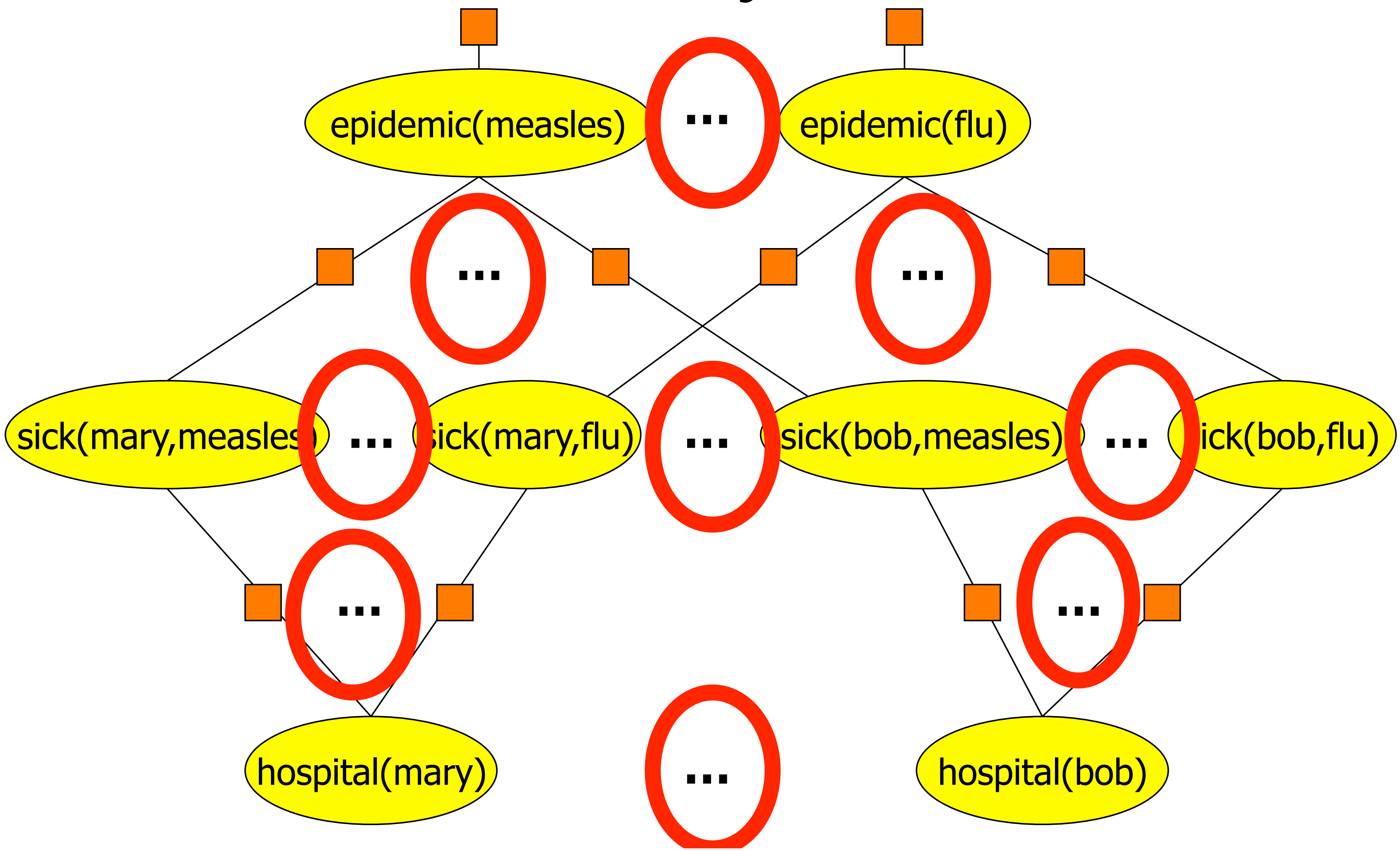
A Lifted inference example

- Dependencies between
  - epidemic(Disease) and sick(Person, Disease)
  - sick(Person, Disease) and hospital(Person)
- Applied to many people and many diseases.

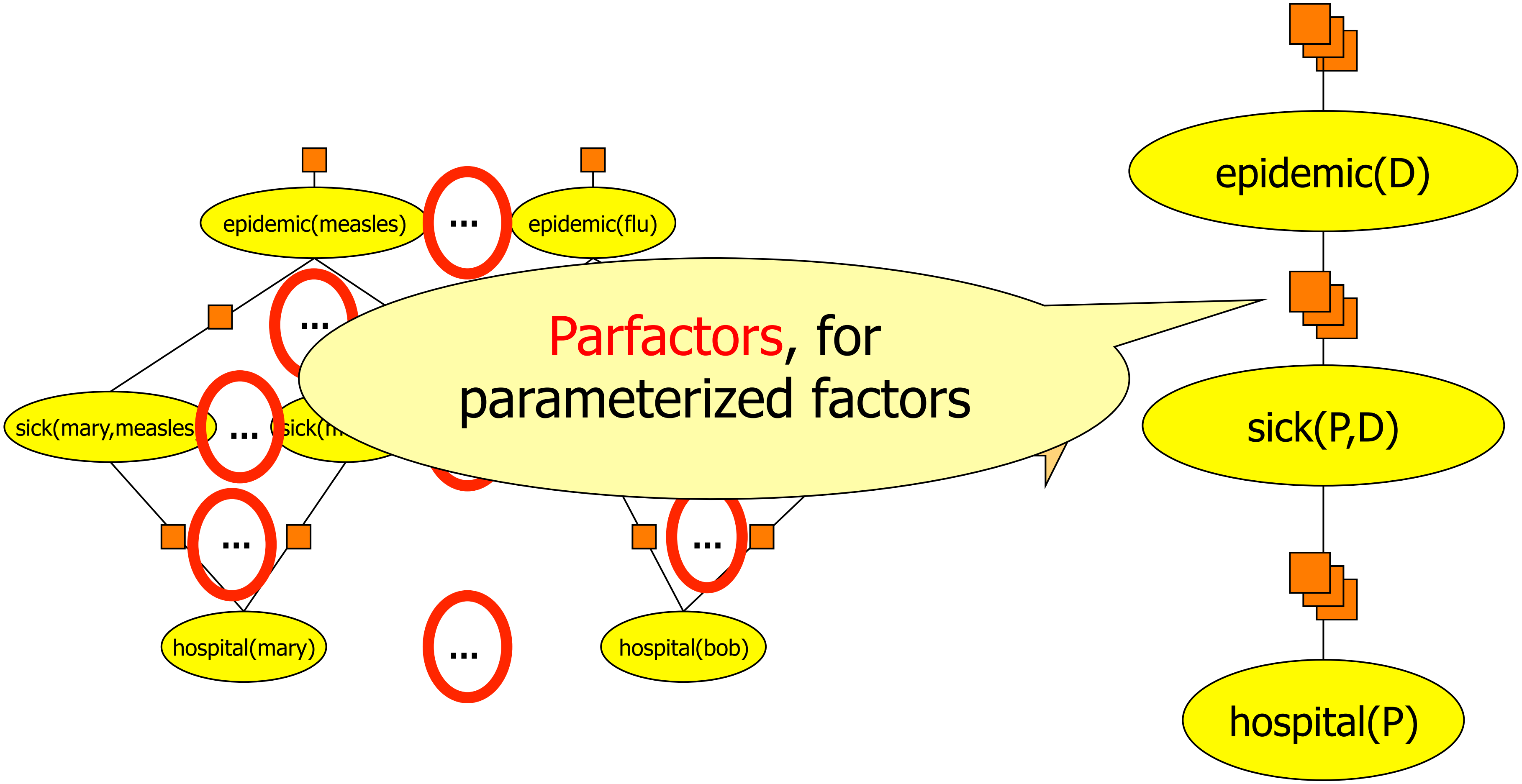
Graphical model size can be large



Much redundancy

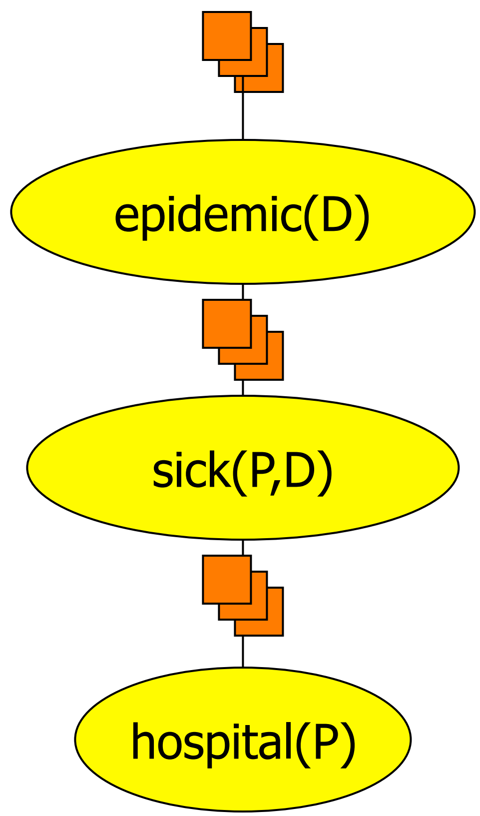


Representing structure



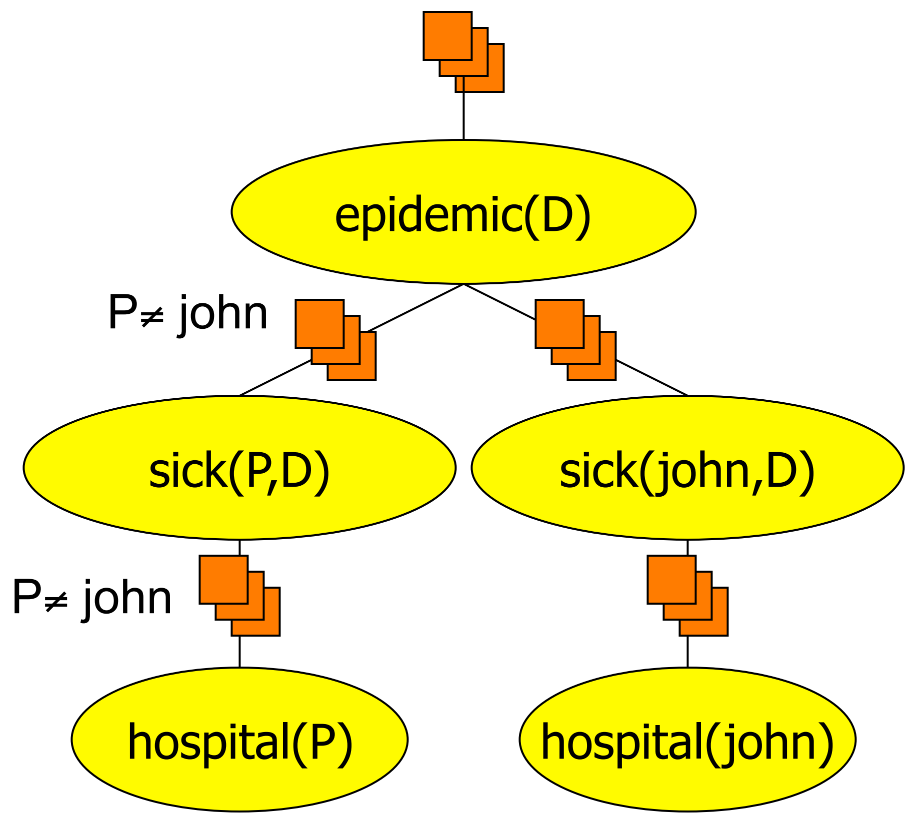
Lifted Inference Example

- P(hospital(john)) = ?



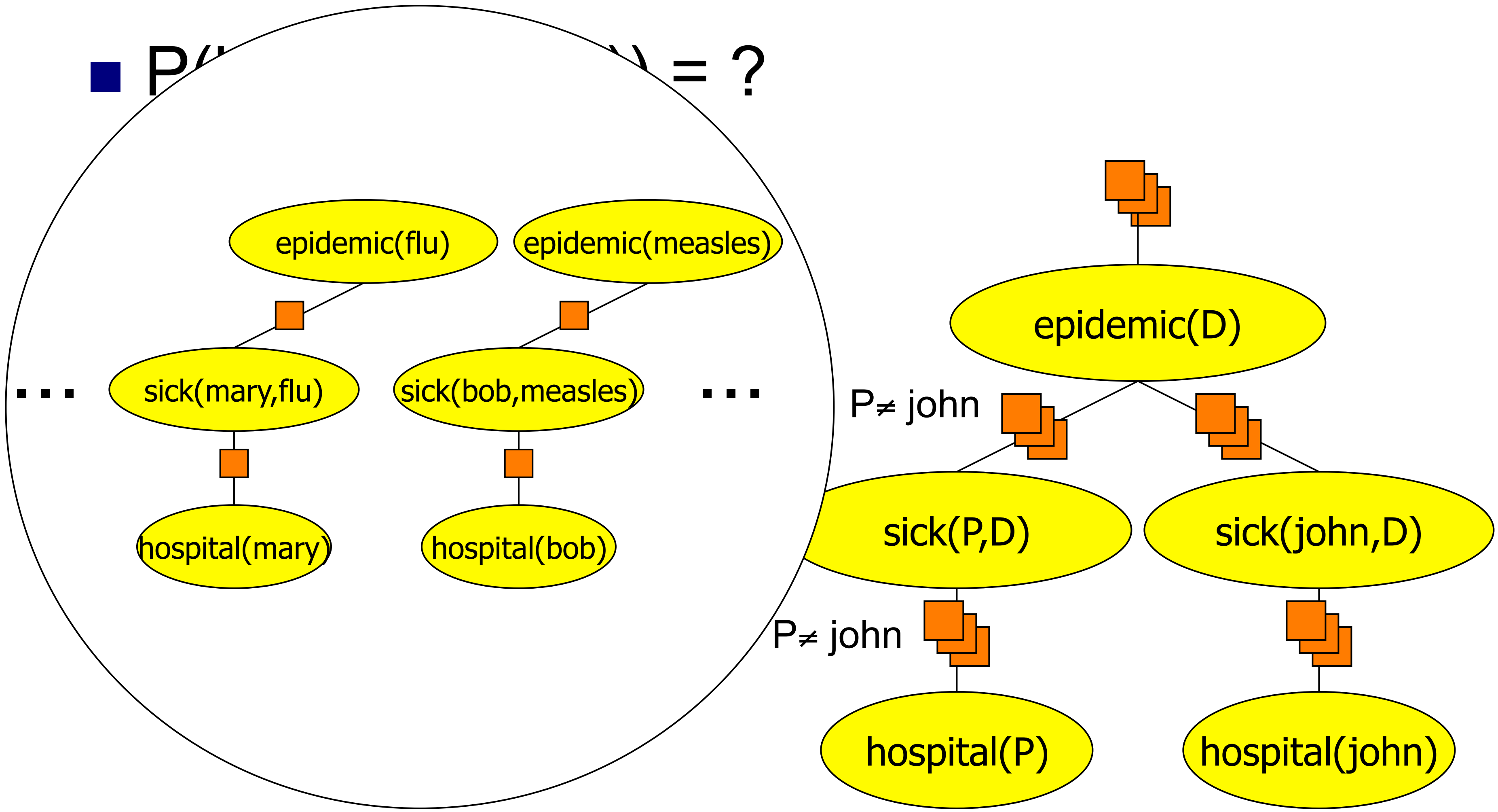
Lifted Inference Example

- P(hospital(john)) = ?



Lifted Inference Example

- P(   ) = ?

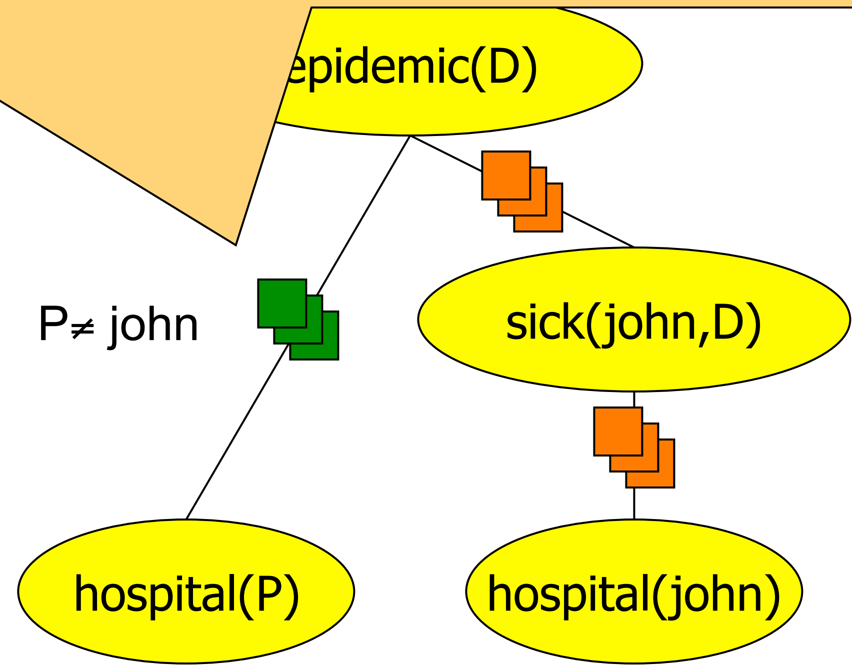


Lifted Inference Example

- P(hospital(john)) = ?

$\phi'(e(D),h(P)) = \sum_{s(P,D):P \neq john} \phi_1(s(P,D),e(D),)\ \phi_2(,s(P,D),h(P))$

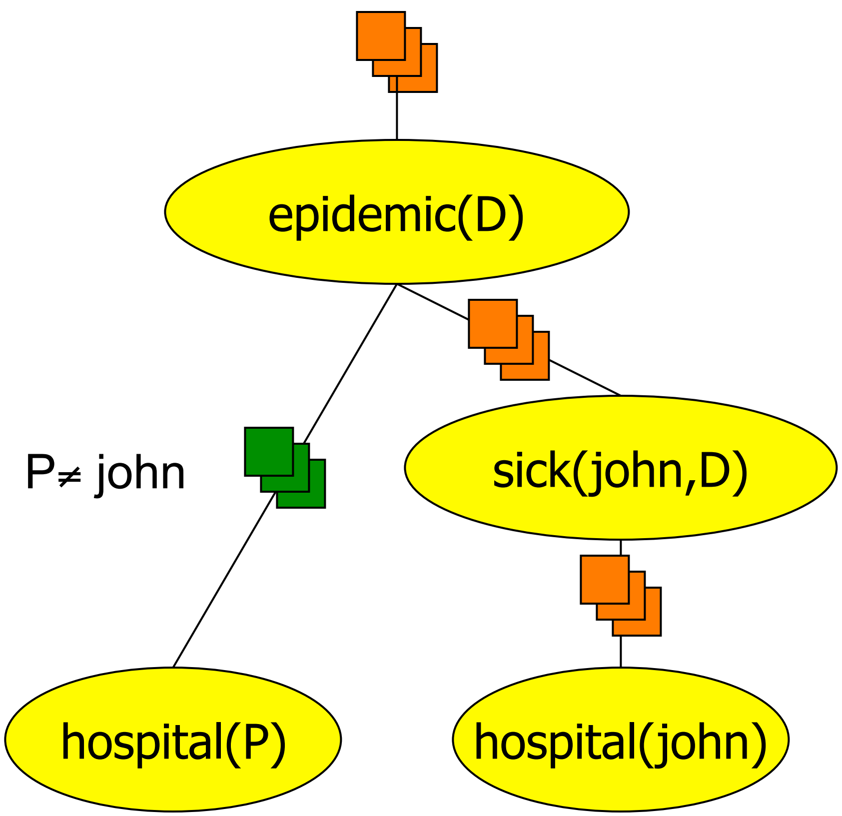
abbreviations



Inversion Elimination

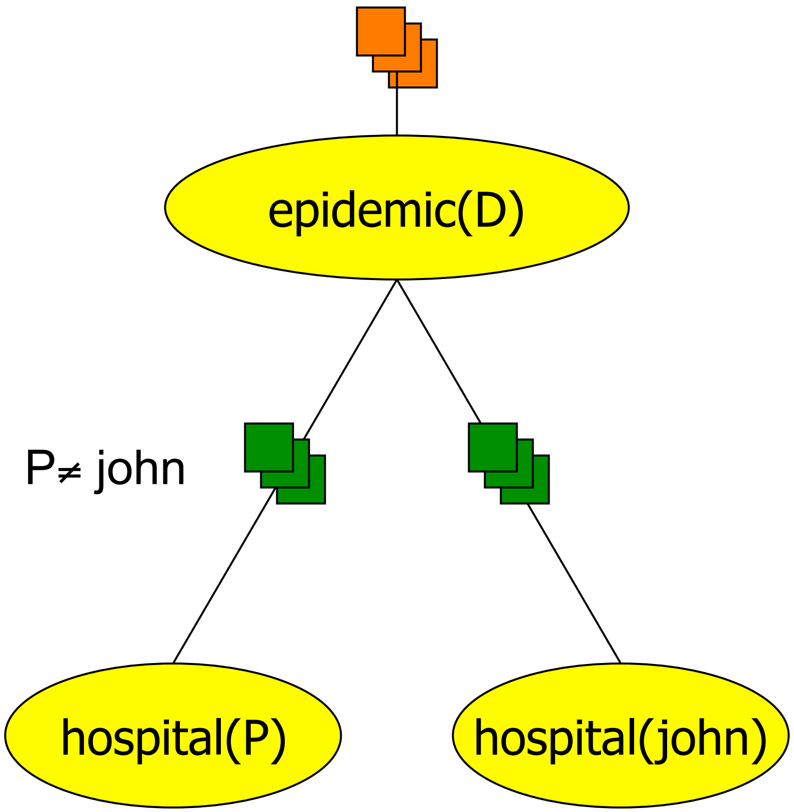
Lifted Inference Example

- P(hospital(john)) = ?



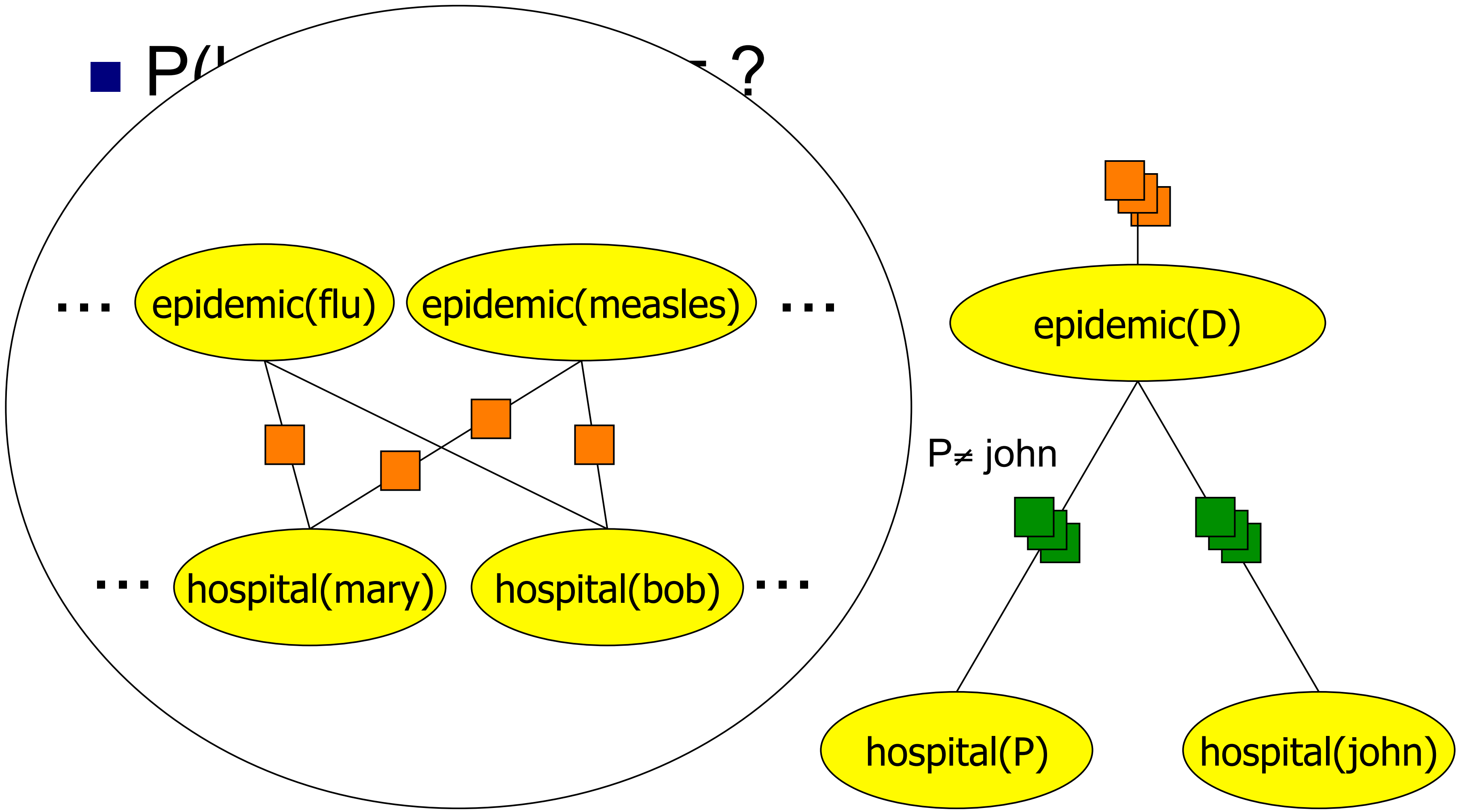
Lifted Inference Example

- P(hospital(john)) = ?



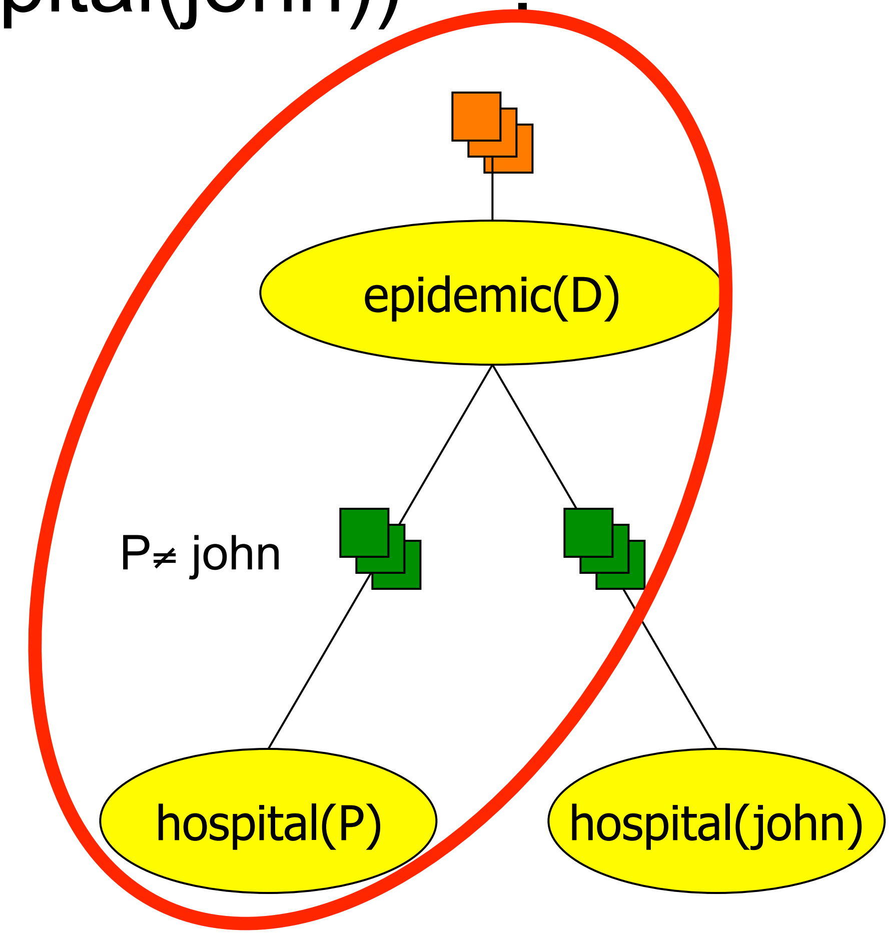
Lifted Inference Example

- P(= ?



Lifted Inference Example

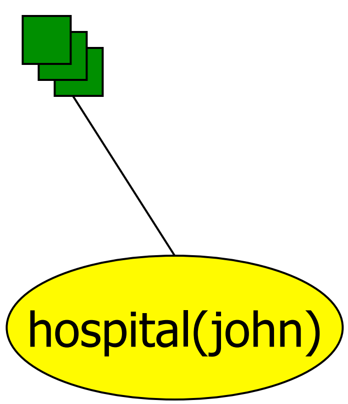
- P(hospital(john)) = ?



Lifted Inference Example

- P(hospital(john)) = ?

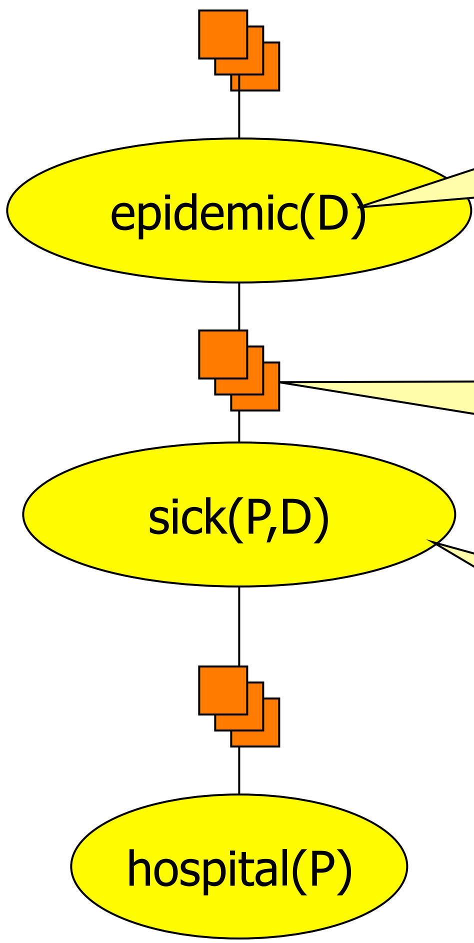
Counting
Elimination



Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion

Representing structure



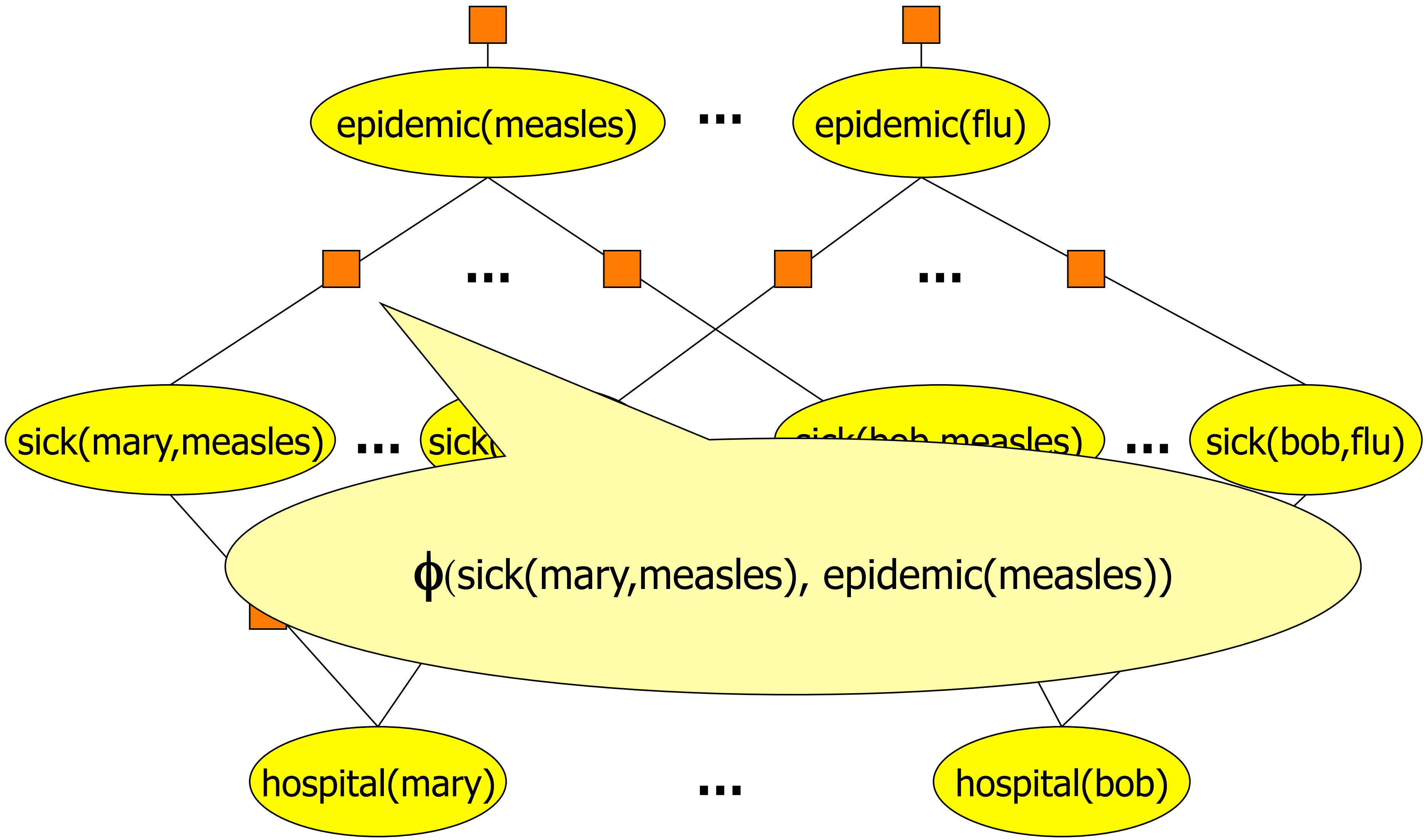
- More comp
- Represents structure explicitly
- n

Logical Variables parameterize random variables

Parfactors, for parameterized factors

Atoms represent a set of random variables

Semantics



epidemic(measles)
…
epidemic(flu)
…
…
sick(mary,measles)
…
…
sick(bob,flu)
ϕ(sick(mary,measles), epidemic(measles))
hospital(mary)
…
hospital(bob)

Making use of structure in Inference

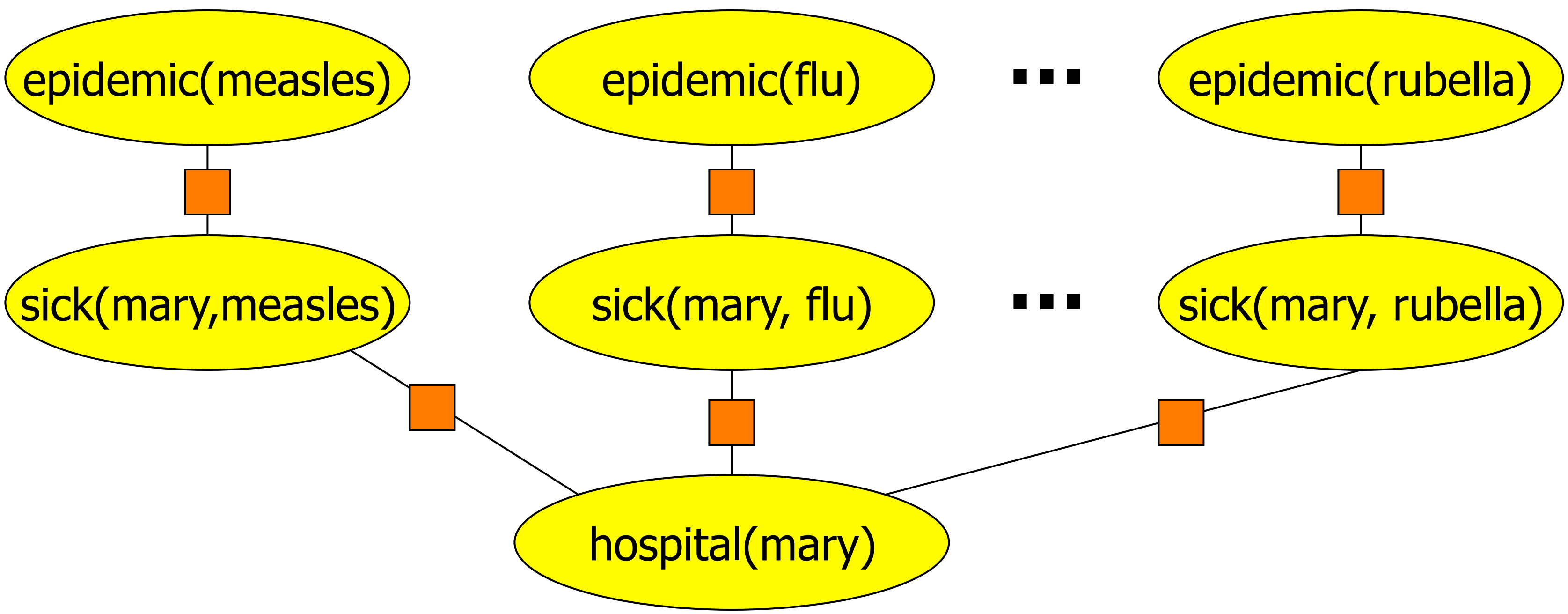
- Task: calculate marginals and posteriors:
  P(sick(bob, measles) | sick(mary,measles)) = ?
- Three approaches
  - plain propositionalization
  - dynamic construction ("smart" propositionalization)
  - lifted inference

Inference - Plain Propositionalization

- Instantiation of potential function for each instantiation
- Lots of redundant computation
- Lots of unnecessary random variables

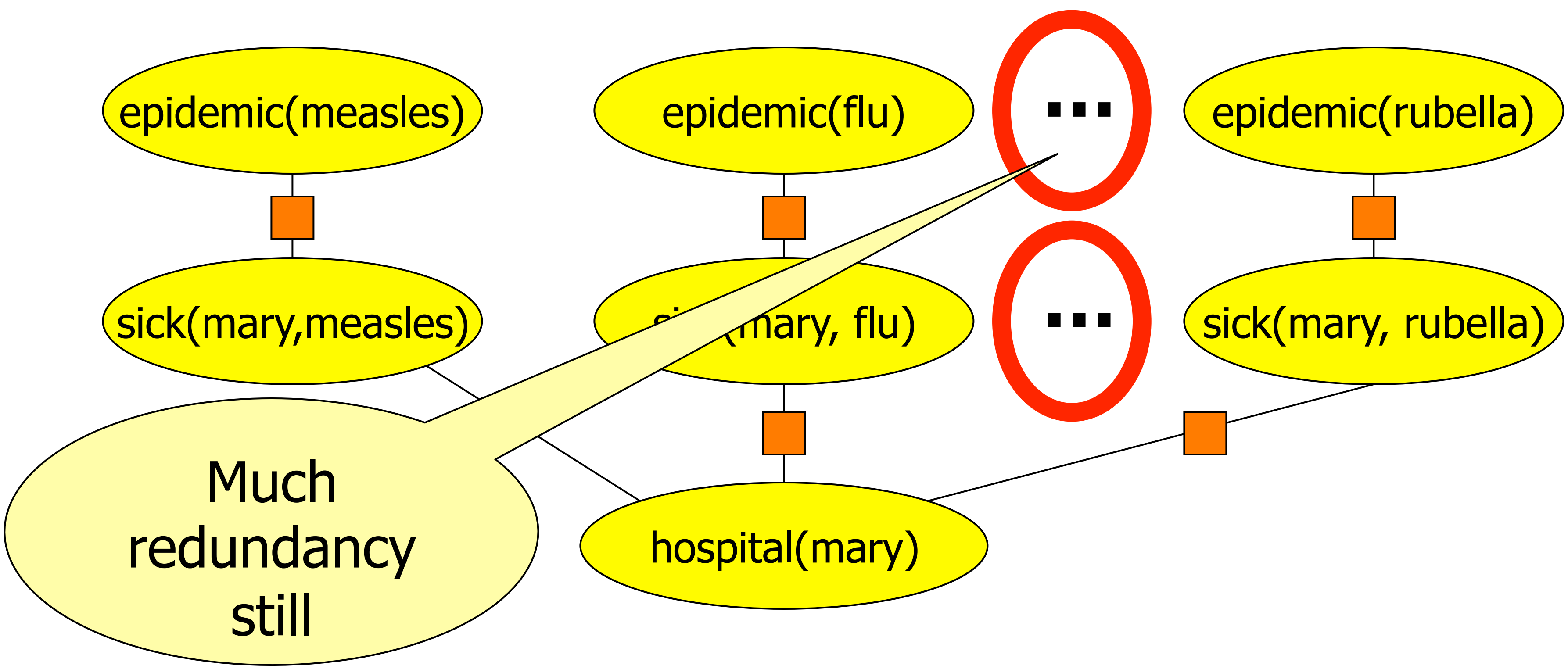
Inference - Dynamic construction

- Instantiation of potential function for relevant parts
- P(hospital(mary) | no epidemics) = ?



Inference - Dynamic construction

- Instantiation of potential function for relevant parts
- P(hospital(mary) | no epidemics) = ?



Inference - Dynamic construction

Used to be the most common approach for exact First-order Probabilistic inference:

- Knowledge-based model construction (Breese, 1992)
- Probabilistic Logic Programming (Ng & Subrahmanian, 1992)
- Probabilistic Logic Programming (Ngo and Haddawy, 1995)
- Probabilistic Relational Models (Friedman et al., 1999)
- Relational Dependency Networks (Neville & Jensen, 2004)
- Relational Bayesian networks (Jaeger, 1997)
- Bayesian Logic Programs (Kersting & DeRaedt, 2001)
- MEBN (Laskey, 2004)
- Markov Logic Networks (Richardson & Domingos, 2004)

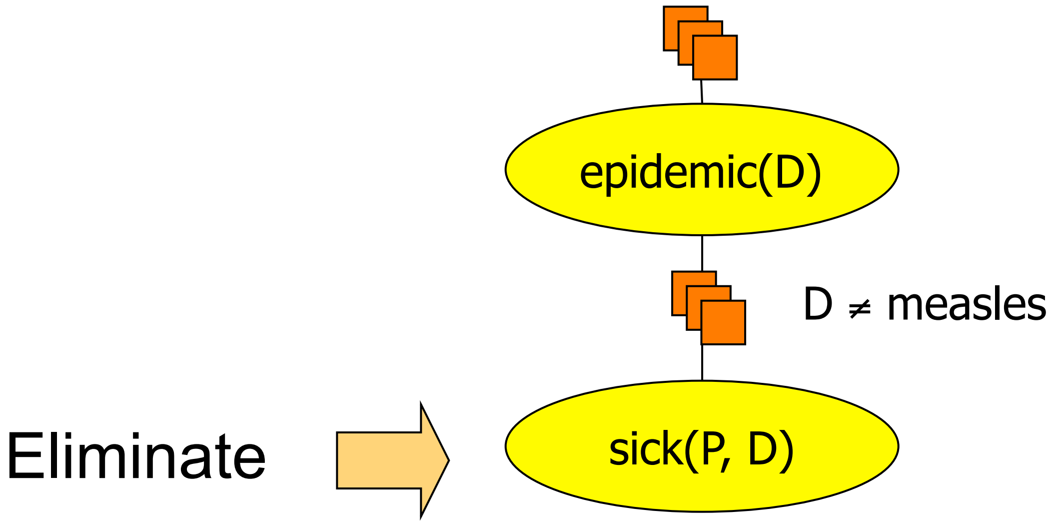
Inference - Lifted inference

- Inference on parameterized, or first-order, level;
- Performs certain inference steps once for a class of random variables;
- Poole (2003) describes a generalized Variable Elimination algorithm which we call Inversion Elimination.
- We formalized Inversion Elimination, showed its limitations and introduced Counting Elimination (IJCAI'05) and Partial Inversion (AAAI'06)

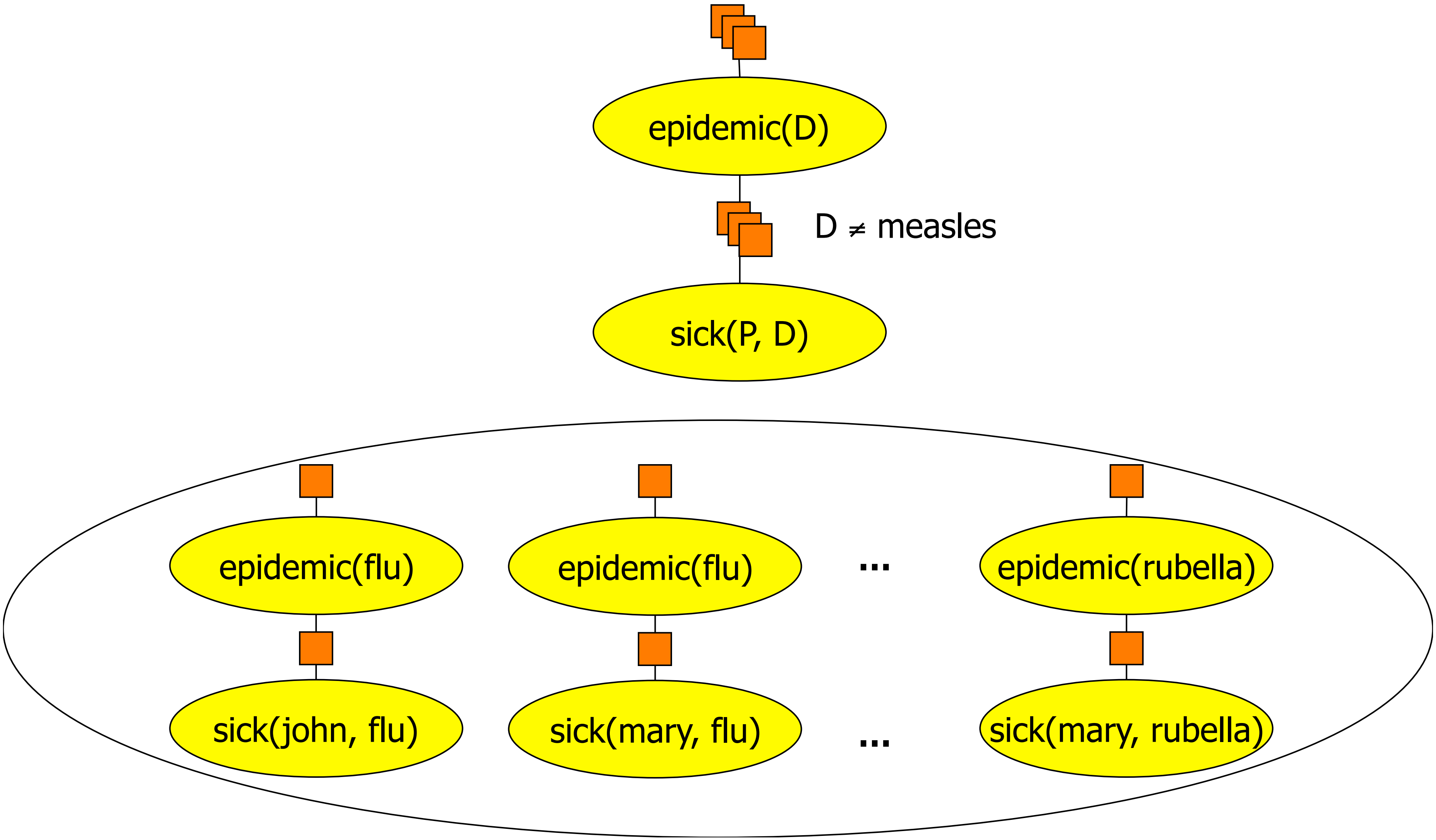
Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion

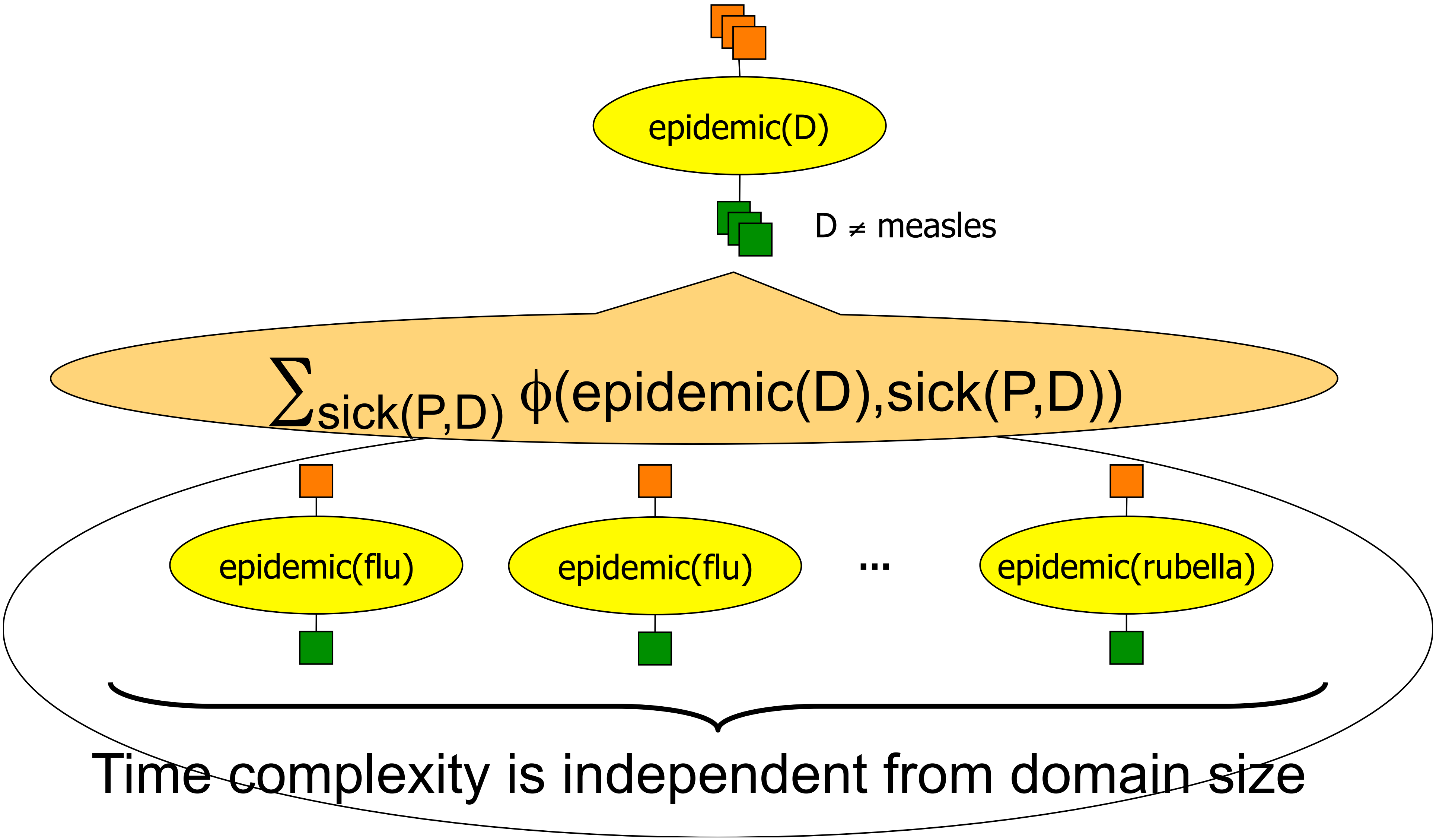
Inversion Elimination (IE)



Inversion Elimination (IE)

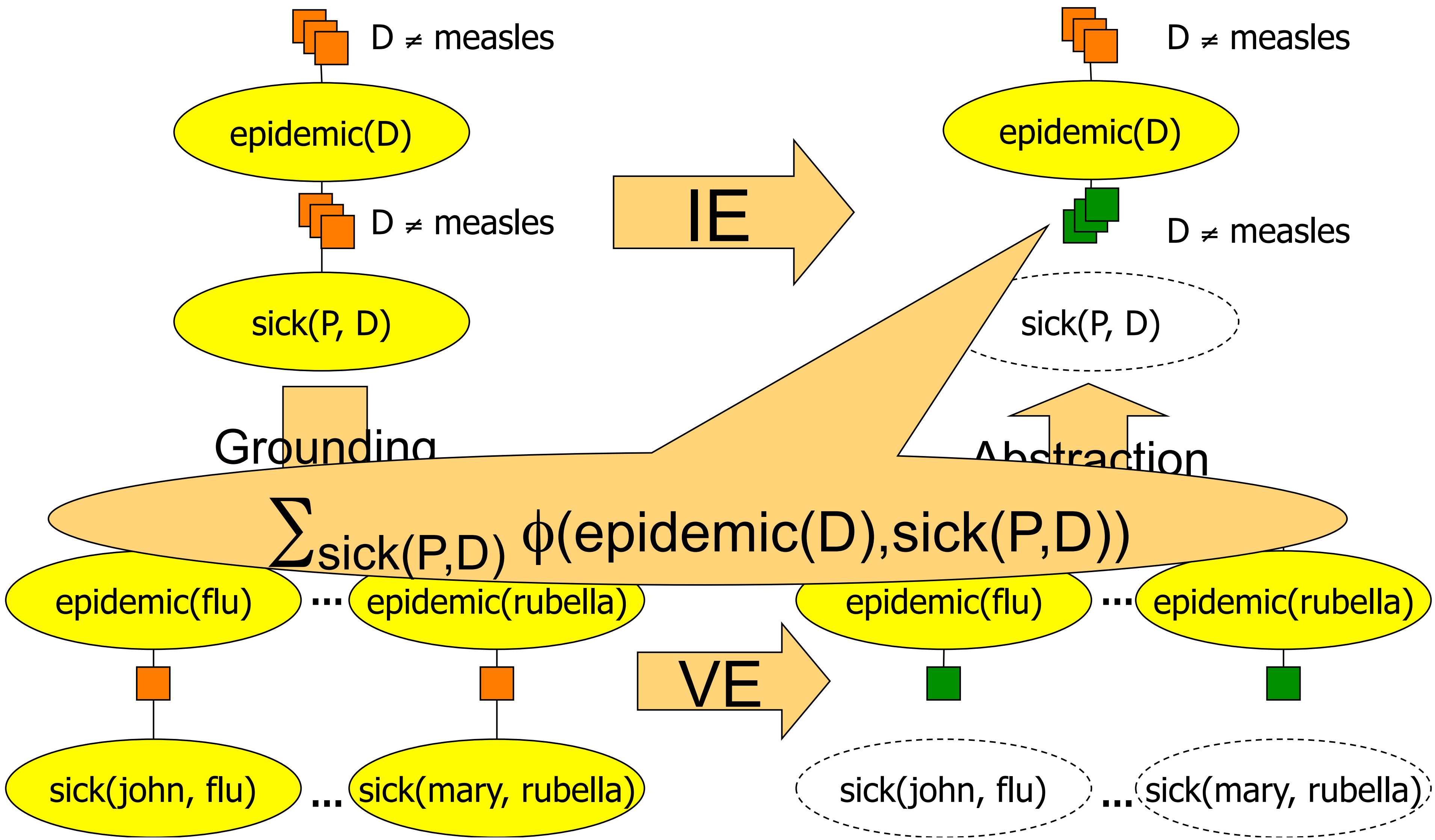


Inversion Elimination (IE)



Time complexity is independent from domain size

Inversion Elimination (IE)



Formalization of IE

- Joint distribution

$$\prod_D \phi_1(e(D)) \quad \prod_{P,D} \phi_2(e(D),s(P,D))$$

- Marginalization by eliminating class s(P,D):

$$\sum_{s(\cdot,\cdot)} \prod_D \phi_1(e(D)) \quad \prod_{P,D} \phi_2(e(D),s(P,D))$$

$$= \prod_D \phi_1(e(D)) \quad \sum_{s(\cdot,\cdot)} \prod_{P,D} \phi_2(e(D),s(P,D))$$

Formalization of IE

$$\sum_{s(\cdot,\cdot)} \prod_{P,D} \phi_2(e(D),s(P,D))$$

$$= \prod_{P,D} \sum_{s(P,D)} \phi_2(e(D),s(P,D))$$

$$= \prod_{P,D} \phi_3(e(D))$$

$$= \prod_{D} \phi_3^{|P|}(e(D))$$

$$= \prod_{D} \phi_4(e(D))$$

Formalization of IE

$$\sum_{s(\cdot,\cdot)} \prod_{P,D} \phi_2(e(D),s(P,D))$$

$$= \sum_{s(p1,d1)} \sum_{s(p1,d2)} \cdots \sum_{s(pn,dm)} \phi_2(e(d_1),s(p_1,d_1))\phi_2(e(d_1),s(p_1,d_2))\ldots \phi_2(e(d_n),s(p_n,d_m))$$

$$= \left(\sum_{s(p1,d1)} \phi_2(e(d_1),s(p_1,d_1))\right)\ldots \left(\sum_{s(p1,d2)} \phi_2(e(d_1),s(p_1,d_2))\right)\ldots \left(\sum_{s(pn,dm)} \phi_2(e(d_n),s(p_n,d_m))\right)$$

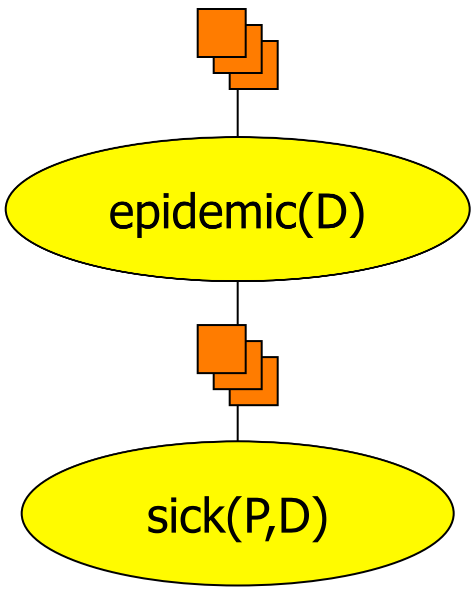
$$= \prod_{P,D} \sum_{s(P,D)} \phi_2(e(D),s(P,D))$$

Formalization of IE

- Inversion Elimination requires certain conditions for its application:
  - Eliminated atom must contain all logical variables in parfactor involved;
  - Eliminated atom instances must not occur together in the same instance of parfactor.
- These are both conditions for parallelism.

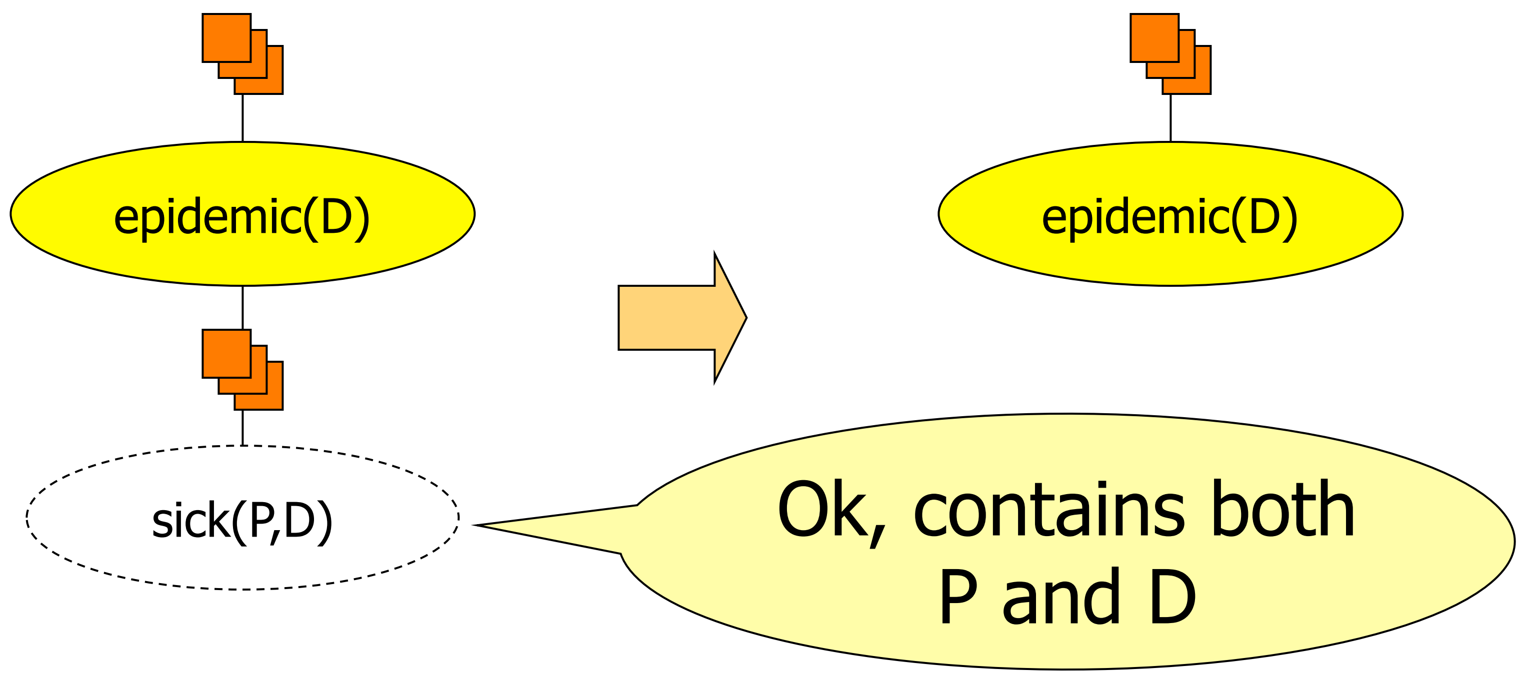
Inversion Elimination - Limitations - I

- Eliminated atom must contain all logical variables in parfactors involved.



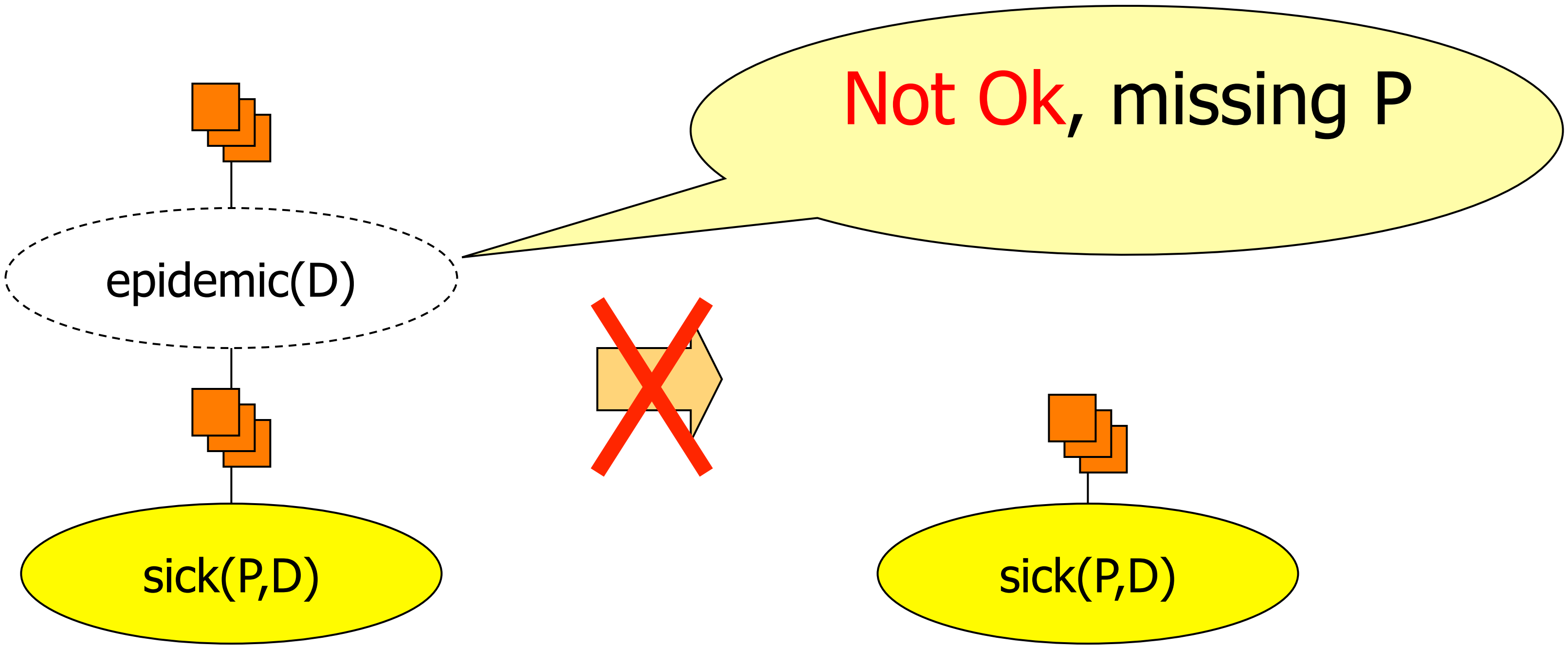
Inversion Elimination - Limitations - I

- Eliminated atom must contain all logical variables in parfactors involved.



Inversion Elimination - Limitations - I

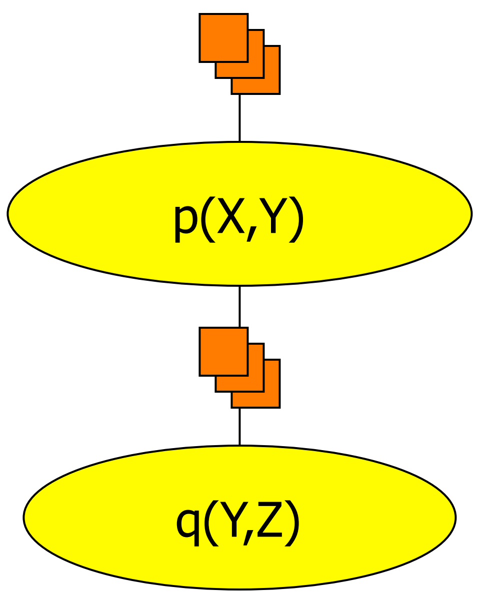
- Eliminated atom must contain all logical variables in parfactors involved.



Inversion Elimination - Limitations - I

- Eliminated atom must contain all logical variables in parfactors involved.

No atom can be eliminated



Inversion Elimination - Limitations - I

- Marginalization by eliminating class e(D):

$$\sum_{e(.)} \prod_{D,P} \phi_2(e(D),s(P,D))$$

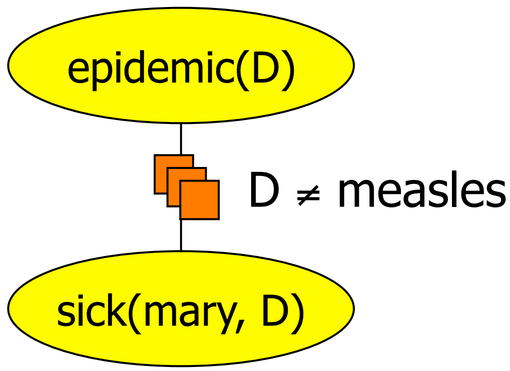
$$= \sum_{e(d1)} \cdots \sum_{e(dn)} \prod_{P} \phi_2(e(d_1),s(P,d_1)) \ldots \prod_{P} \phi_2(e(d_n),s(P,d_n))$$

$$= \left(\sum_{e(d1)} \prod_{P} \phi_2(e(d_1),s(P,d_1))\right) \ldots \left(\sum_{e(dn)} \prod_{P} \phi_2(e(d_n),s(P,d_n))\right)$$

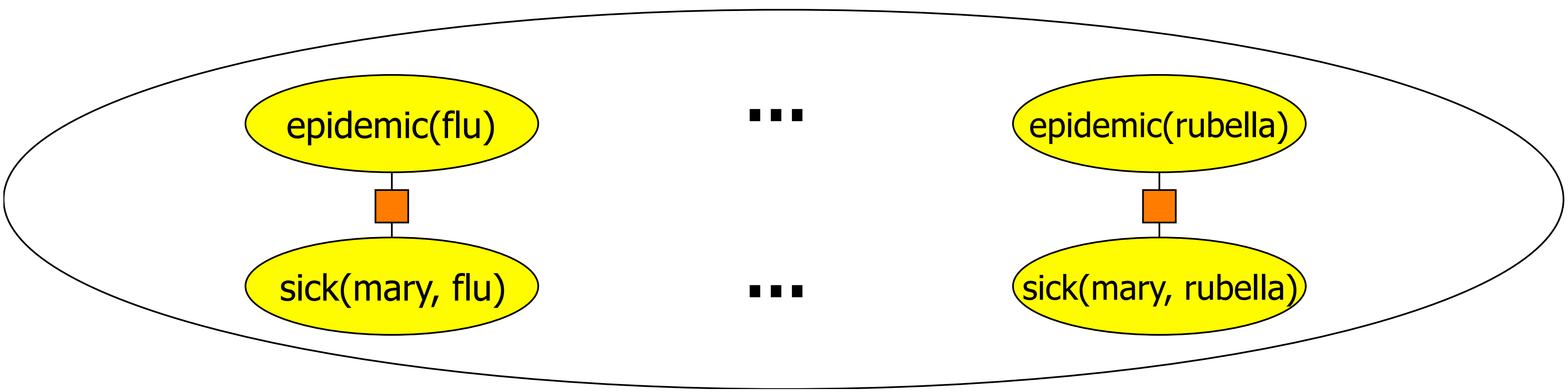
$$= \prod_{D} \sum_{e(D)} \prod_{P} \phi_2(e(D),s(P,D))$$

Inversion Elimination - Limitations - II

- Requires eliminated RVs to occur in separate instances of parfactor

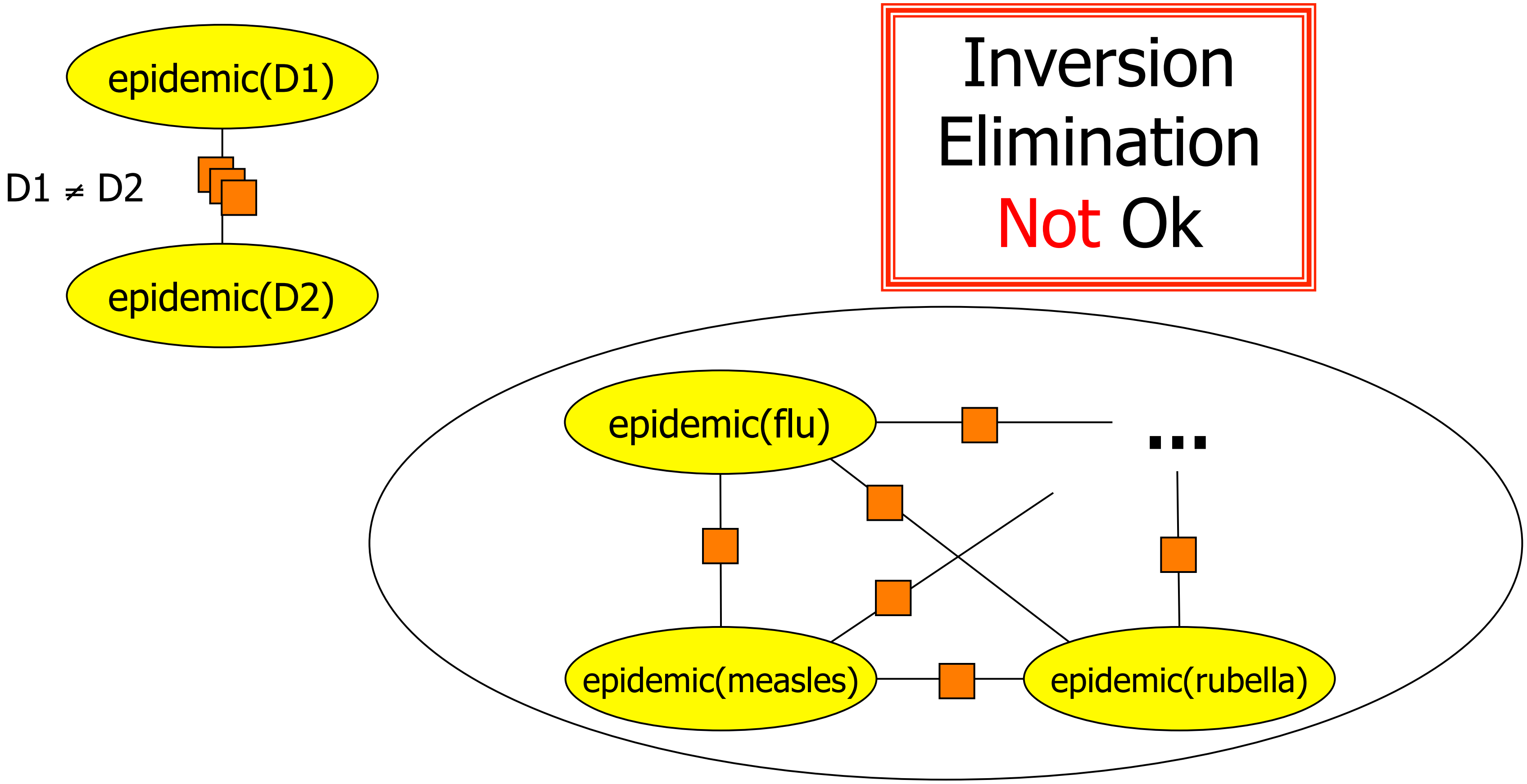


Inversion Elimination Ok



Inversion Elimination - Limitations - II

- Requires eliminated RVs to occur in separate instances of parfactor



Inversion Elimination - Limitations - II

$$\sum_{e(\cdot)} \prod_{D1 \neq D2} \phi(e(D_1),e(D_2))$$

$$= \sum_{e(d1)} \cdots \sum_{e(dn)} \phi(e(d_1),\ e(d_2)) \ldots \phi(e(d_{n-1}),e(d_n))$$

$$= \sum_{e(d1)} \phi(e(d_1),\ e(d_2)) \ldots \sum_{e(dn)} \phi(e(d_{n-1}),e(d_n))$$

Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion

Counting Elimination

$\sum_{e(\cdot)} \prod_{D1 \neq D2}$ $\phi(e(D$ ...

$= \sum_{e(\cdot)}$ ϕ ...

ϕ ...

$\phi(1,$ ...

$\phi(1,1)$ ... e(.),D1≠D2

$= \sum_{e(.)} \prod_{v} \phi(v)^{\#v \text{ in } e(.),D1 \neq D2}$

$= \sum_{i=0}^{|e(.)|} \binom{|e(.)|}{i} \prod_{v} \phi(v)^{\#v \text{ in } e(.),D1 \neq D2 \text{ (from } i)}$

... es

Does depend on domain size, but not exponentially

#(0,1) = 70 * 30
#(1,1) = 30 * (30 – 1)

Counting Elimination Conditions

- No shared logical variables between atoms, so counting can be done independently

ϕ(epidemic(D_1, Region), epidemic(D_2, Region)) is not suitable for Counting elimination.

Outline

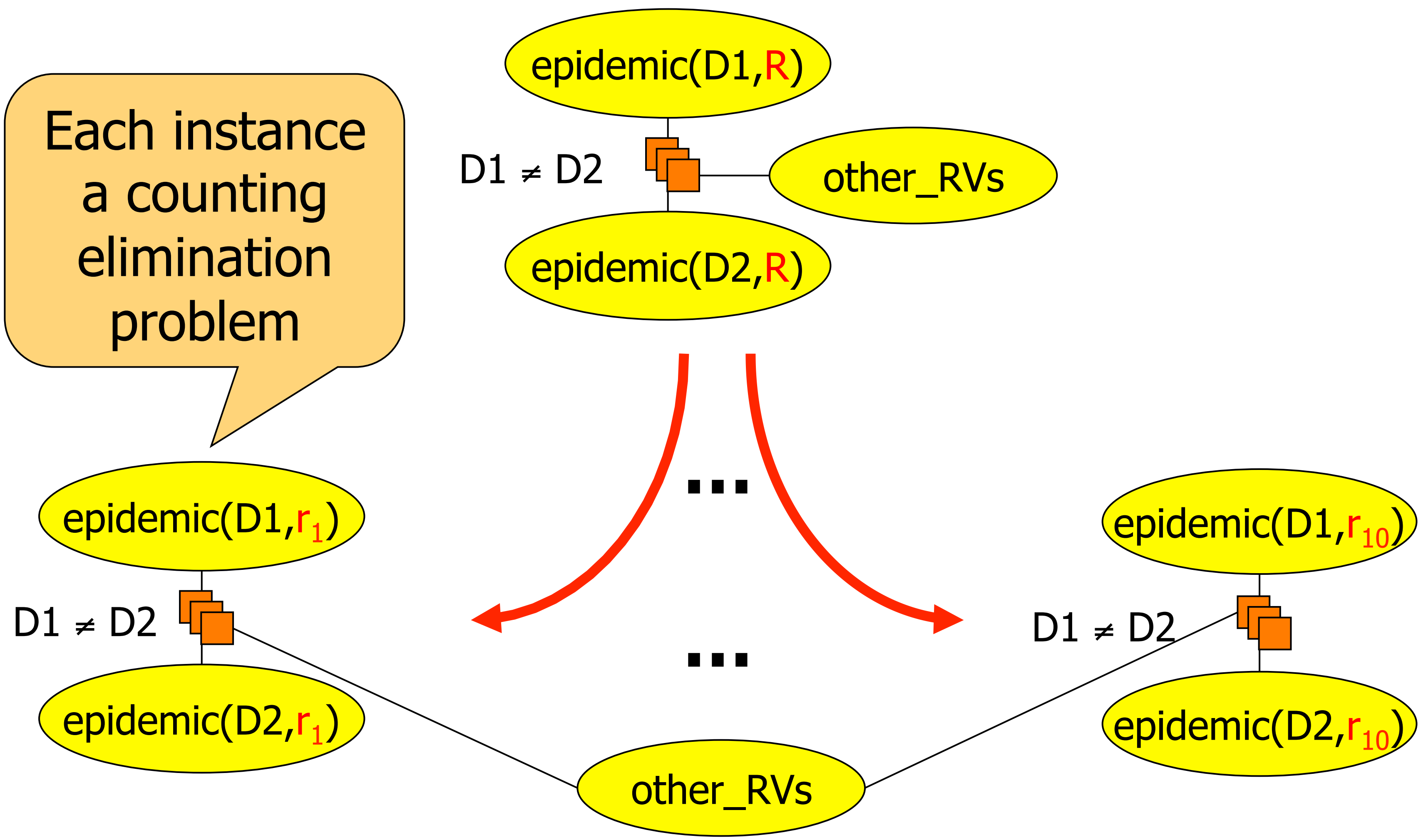
- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion

Some things not covered by both Inversion and Counting

- Eliminating epidemic from

  ϕ(epidemic(Disease1,Region),
  epidemic(Disease2,Region),
  other_RVs)
- No atom with all logical variables
) no Inversion Elimination.
- Shared logical variables
) no Counting Elimination.

Partial Inversion, graphically



Partial Inversion

$$\sum_{e(.,.)} \prod_{D1 \neq D2, R} \phi(\, e(D_1,R),\, e(D_2,R),\, o\,)$$

$$= \prod_{R} \sum_{e(.,R)} \prod_{D1 \neq D2} \phi(\, e(D_1,R),\, e(D_2,R),\, o\,)$$

$$= \prod_{R} \phi'(\, o\,) = \phi'(\, o\,)^{|R|} = \phi''(\, o\,)$$

R is bound
) no shared logical variables
) Counting Elimination

Generalizes both Inversion and Counting Elimination

Partial Inversion Conditions

ϕ(friends(X,Y), friends(Y,X), sick(X,D), sick(Y,D))

- Cannot partially invert on X,Y because friends(bob,mary) appears in more than one grounding of parfactor.
- But with Partial Inversion we do not need all logical variables in order to invert.

Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion

Conclusion

- Expressive probabilistic representations used *during* inference as well.
- Lifted inference equivalent to grounded inference; much faster, yet yielding the same exact answer.

Thanks!