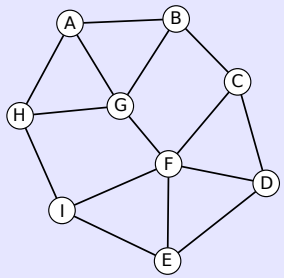
Protein Hypernetworks

Johannes Köster, Eli Zamir, Sven Rahmann

August 20, 2012

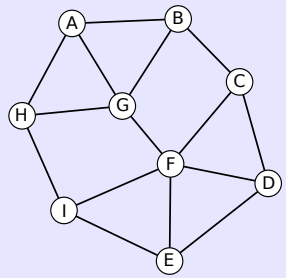
Protein Network Modelling

Interaction maps (undirected graphs)



Protein Network Modelling

Interaction maps (undirected graphs)

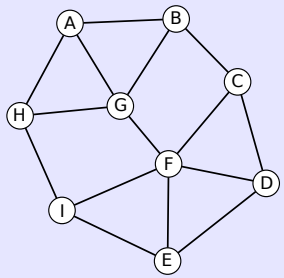


Differential equations (Law of Mass Action), Bayesian Networks, ...

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_{\mathrm{on}}[A][B] - k_{\mathrm{off}}[C]$$

Protein Network Modelling

Interaction maps (undirected graphs)



Differential equations (Law of Mass Action), Bayesian Networks, ...

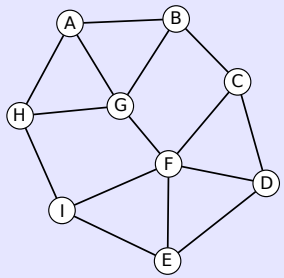
$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_{\mathrm{on}}[A][B] - k_{\mathrm{off}}[C]$$

accuracy ↓

scalability ↑

Protein Network Modelling

Interaction maps (undirected graphs)



Protein Hypernetworks ?

Differential equations (Law of Mass Action), Bayesian Networks, ...

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_{\mathrm{on}}[A][B] - k_{\mathrm{off}}[C]$$

accuracy ↓

scalability ↑

Structure

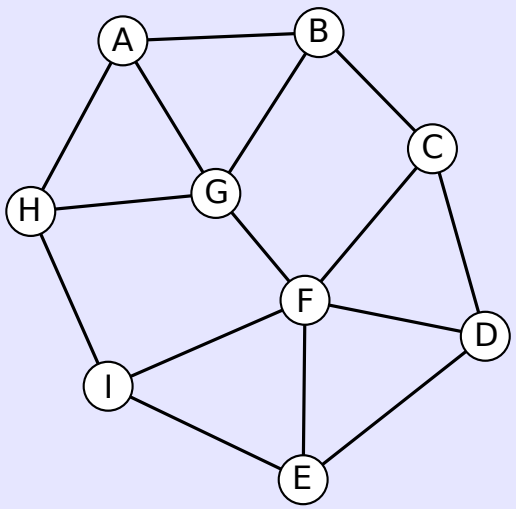
1 Protein Hypernetworks

2 Mining Protein Hypernetworks

3 Data Aquisition

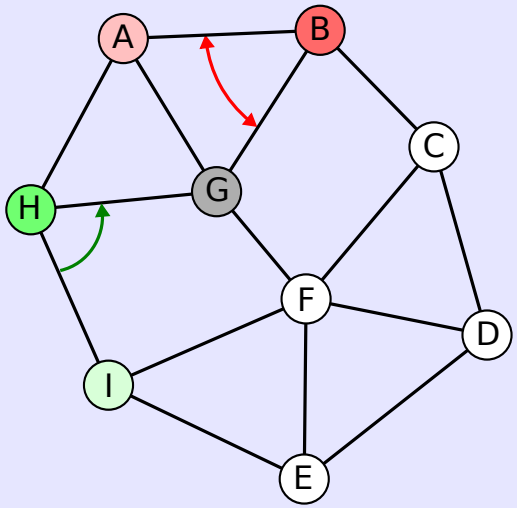
Idea

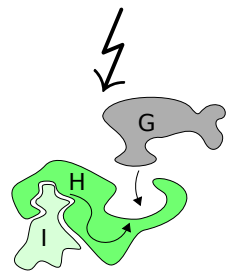
Protein Network (P, I)

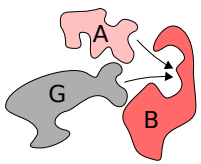


Idea

Protein Network (P, I)



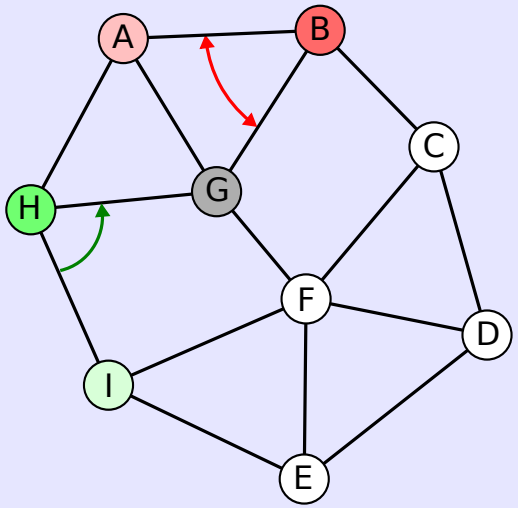


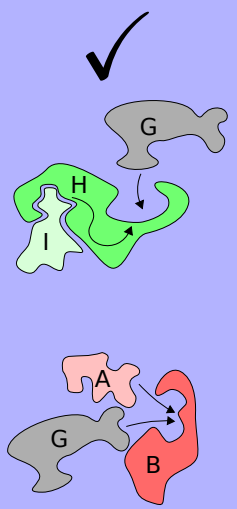


Idea

Protein Hypernetwork (P, I, C)

Protein Network (P, I)





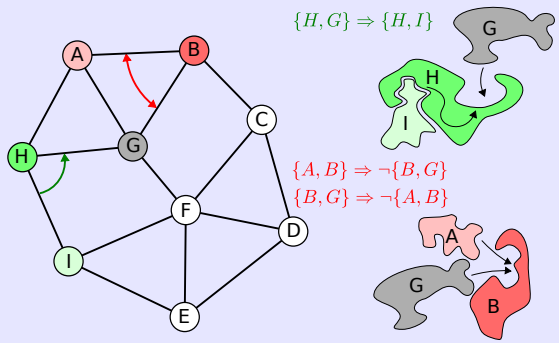
Boolean Logic Constraints C

$$\{G, H\} \Rightarrow \{I, H\}$$

$$\{A, B\} \Rightarrow \neg\{G, B\}$$
$$\{G, B\} \Rightarrow \neg\{A, B\}$$

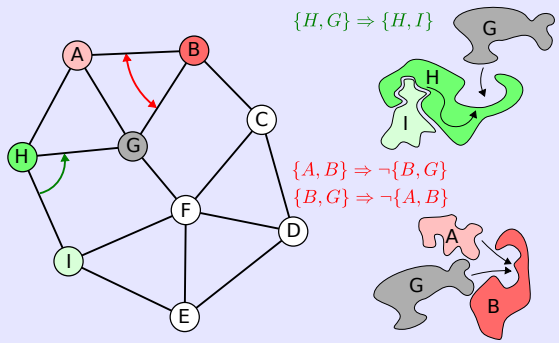
Mining Protein Hypernetworks

Protein Hypernetwork (P, I, C)



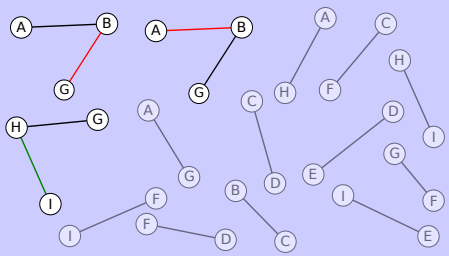
Mining Protein Hypernetworks

Protein Hypernetwork (P, I, C)



Minimal network states (Nec, Imp) for $q \in P \cup I$

$$q \wedge \bigwedge_{c \in C} c$$



- Satisfying model $\alpha : P \cup I \to \{0, 1\}$ by tableau algorithm

$$Nec := \{q' \in P \cup I \mid \alpha(q') = 1\}$$
$$Imp := \{q' \in P \cup I \mid \alpha(q') = 0 \text{ due to active } c \in C\}$$

Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

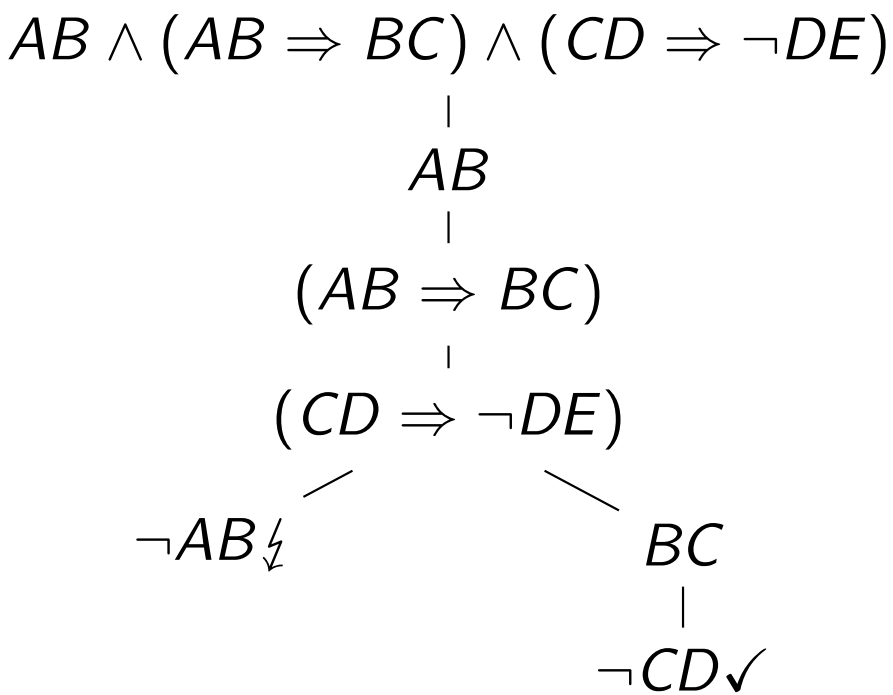


Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model

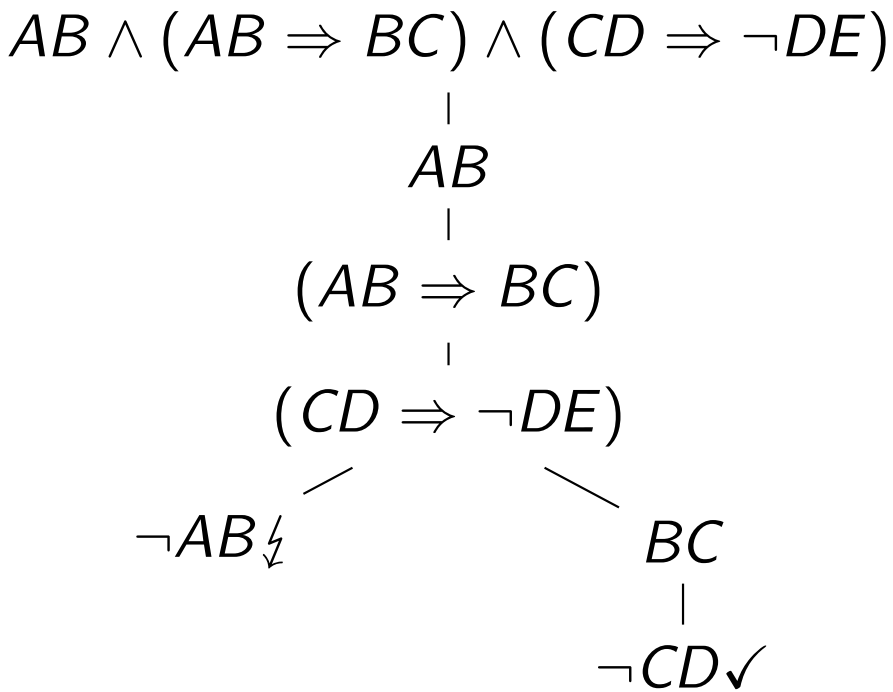


Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model

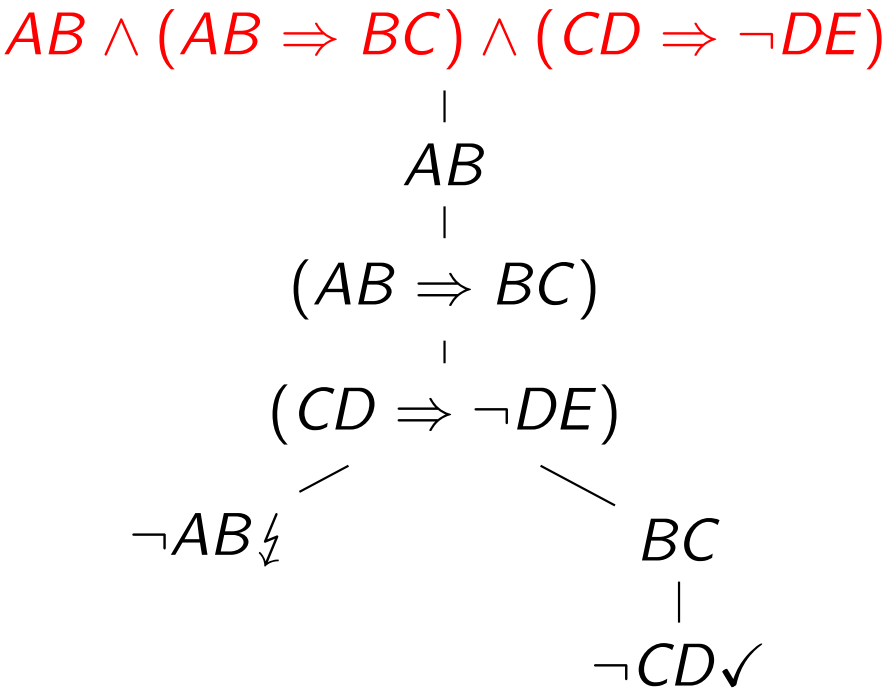


Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model

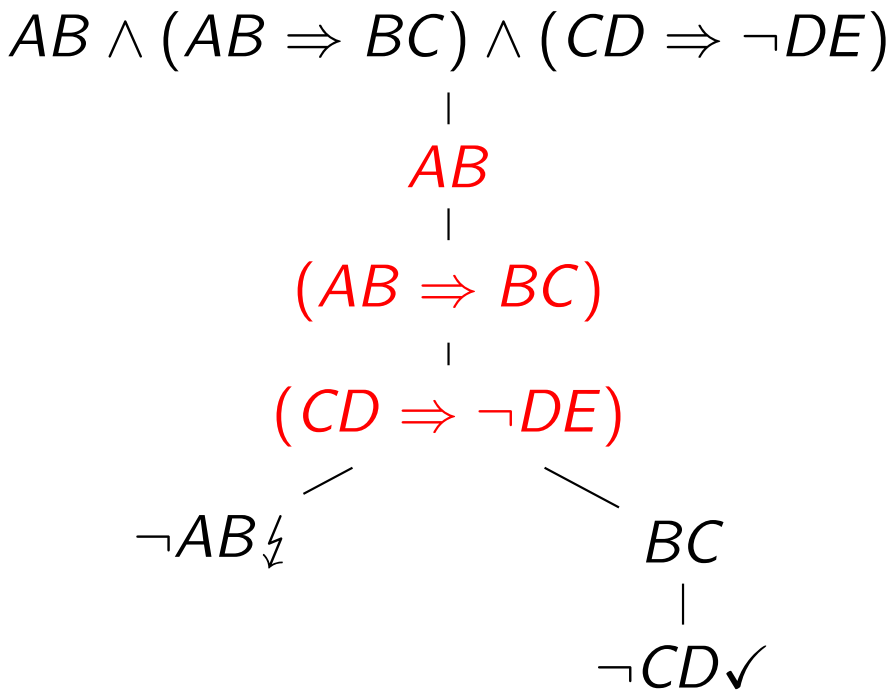


Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model

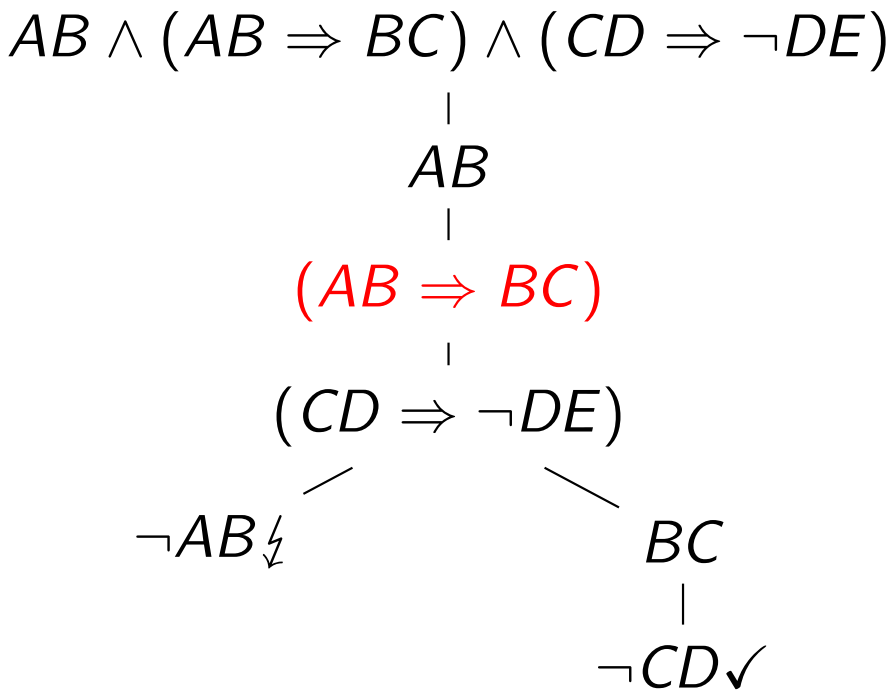


Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model

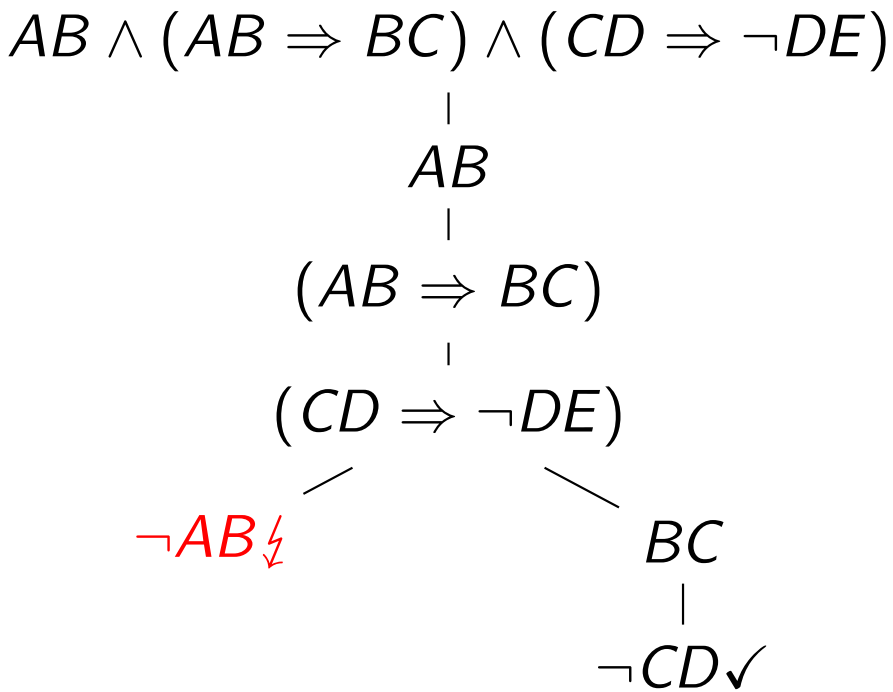


Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model

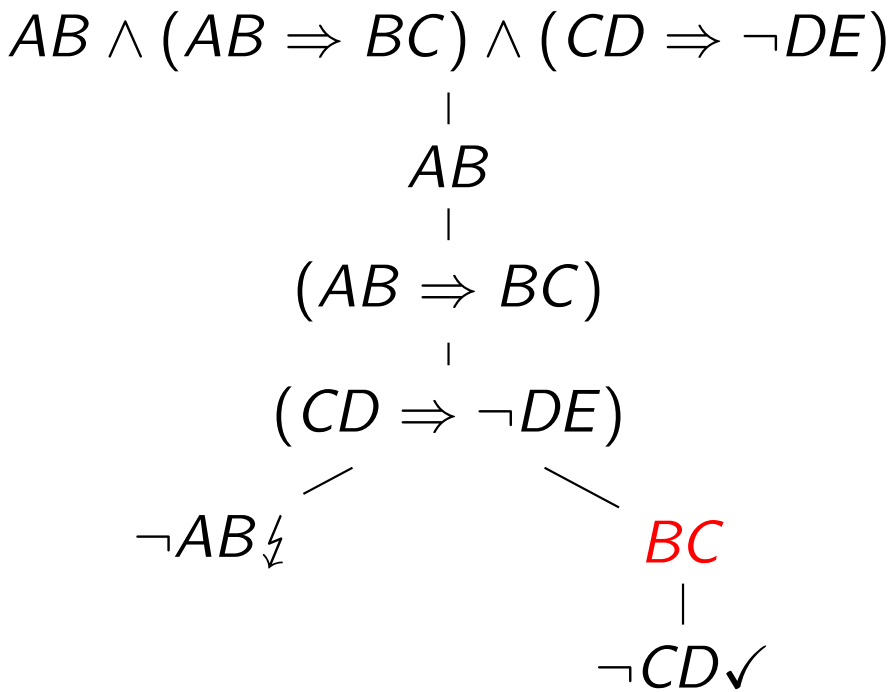


Tableau Algorithm

Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model

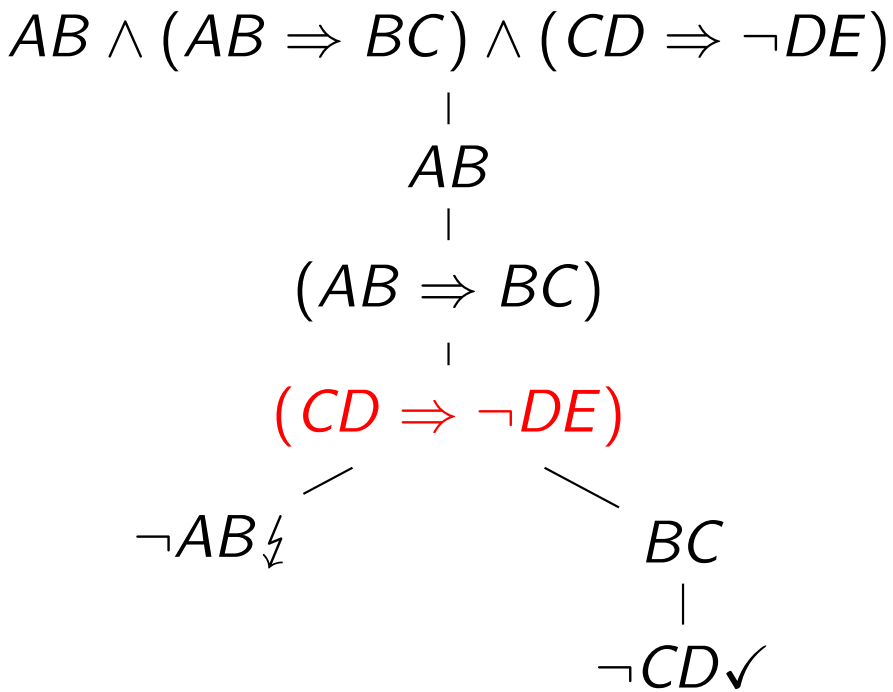


Tableau Algorithm

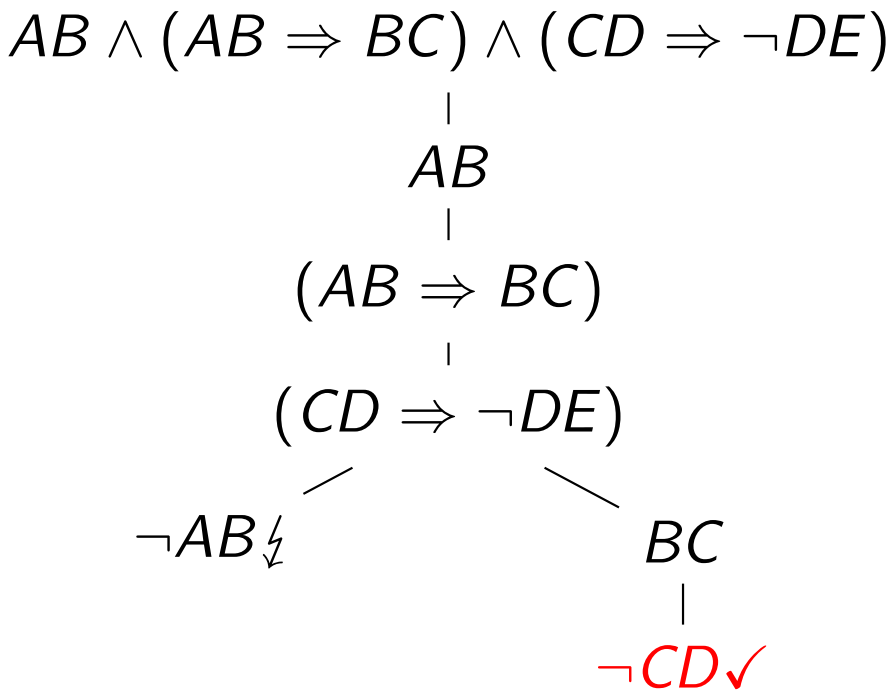
Propositional Logic Tableau Algorithm

for a given formula f

- explore depth-first the tree of deductions from root f
- each root-leaf-path without contradiction is a satisfying model

Modifications

- expand disjunctions from left to right
- allow to pre-block subformulas to guide the algorithm to the right model



Minimal Network States

Clashes

Two minimal network states (Nec, Imp) and (Nec', Imp') are clashing iff

$$Nec \cap Imp' \neq \emptyset \text{ or } Nec' \cap Imp \neq \emptyset.$$

not clashing pair $\rightarrow$ interactions simultaneously possible

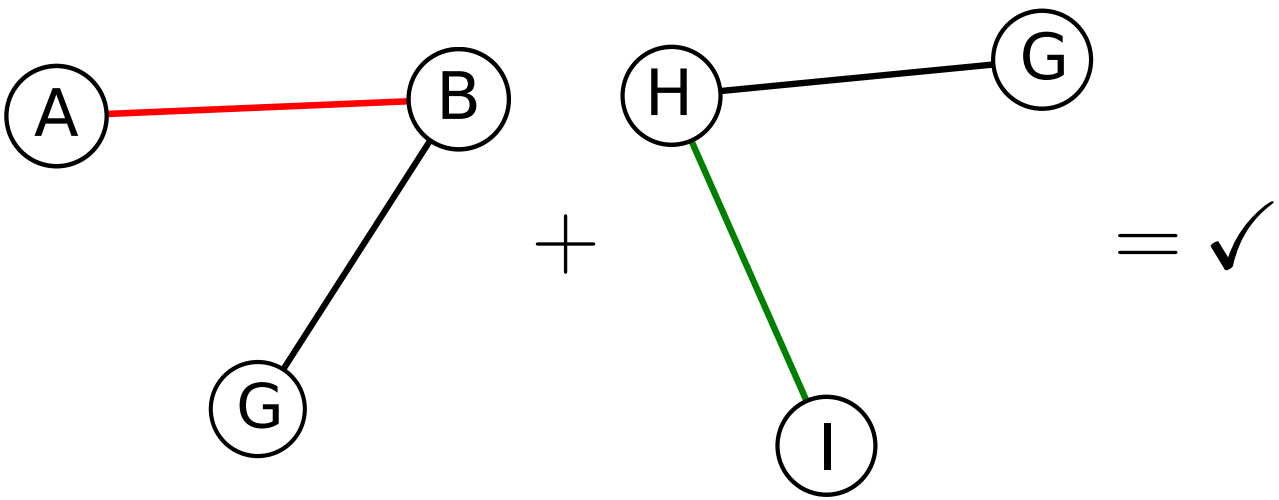
Minimal Network States

Clashes

Two minimal network states (Nec, Imp) and (Nec', Imp') are clashing iff

$$Nec \cap Imp' \neq \emptyset \text{ or } Nec' \cap Imp \neq \emptyset.$$

not clashing pair $\rightarrow$ interactions simultaneously possible



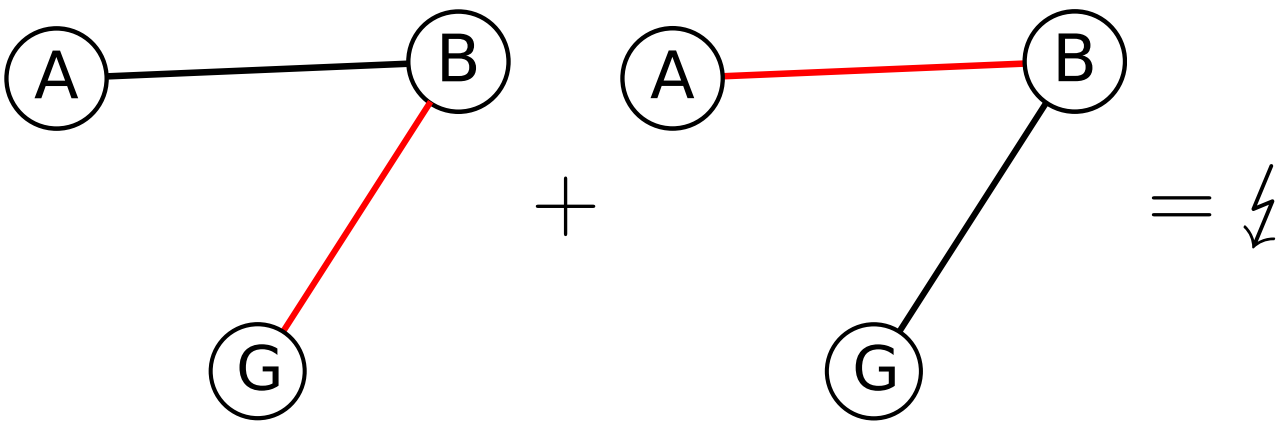
Minimal Network States

Clashes

Two minimal network states (Nec, Imp) and (Nec', Imp') are clashing iff

$$Nec \cap Imp' \neq \emptyset \text{ or } Nec' \cap Imp \neq \emptyset.$$

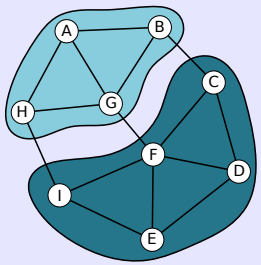
not clashing pair $\rightarrow$ interactions simultaneously possible



Prediction of Protein Complexes

Network based complex prediction

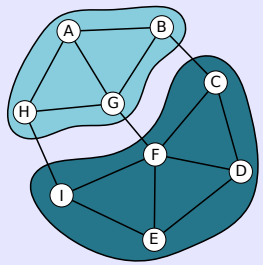
- e.g. dense regions



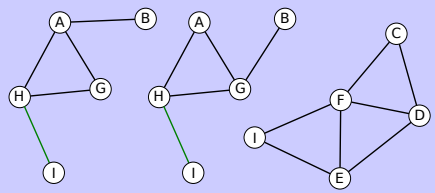
Prediction of Protein Complexes

Network based complex prediction

- e.g. dense regions



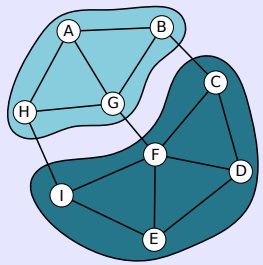
Maximal combinations of minimal network states



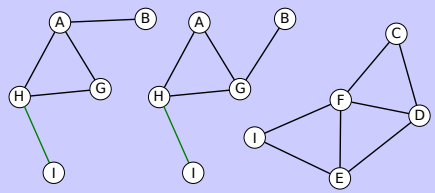
Prediction of Protein Complexes

Network based complex prediction

- e.g. dense regions

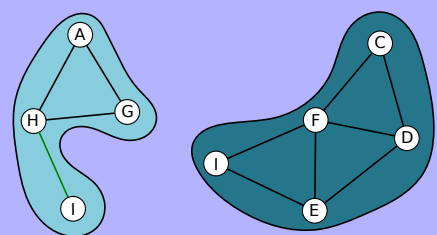


Maximal combinations of minimal network states



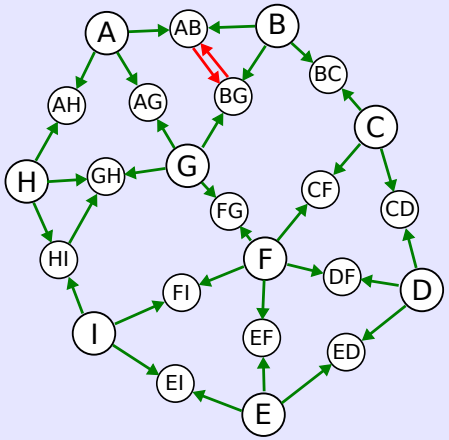
Refined complexes

- no violated constraints

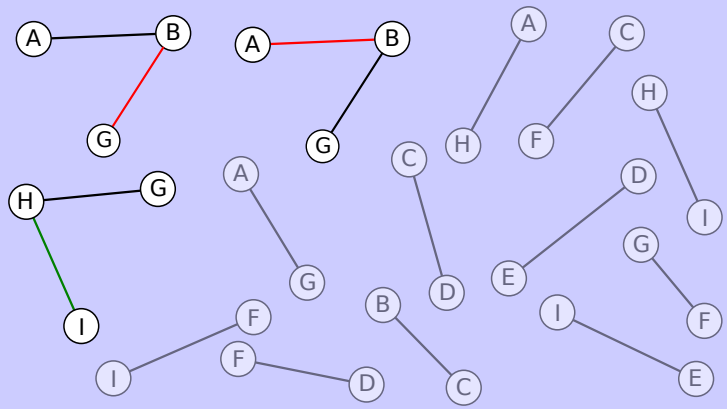


Prediction of Functional Importance

Minimal network state graph

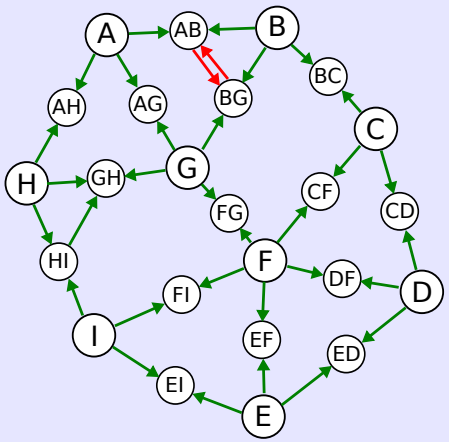


Minimal network states

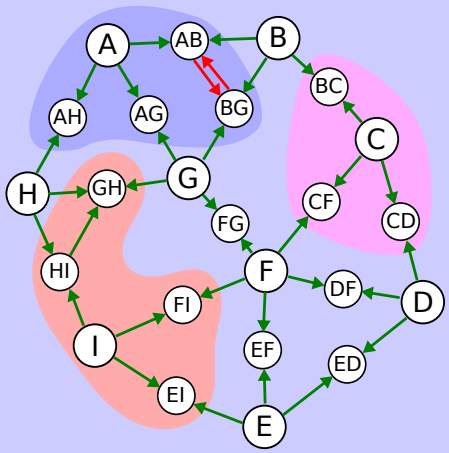


Prediction of Functional Importance

Minimal network state graph

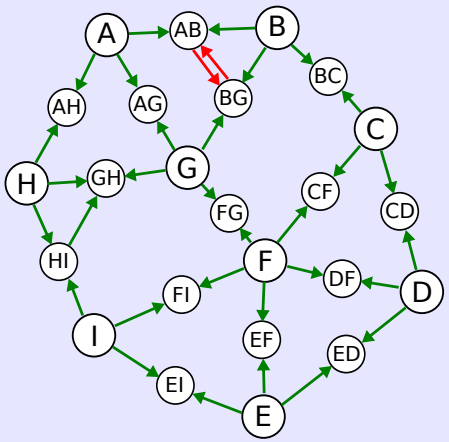


Breadth first search from each node

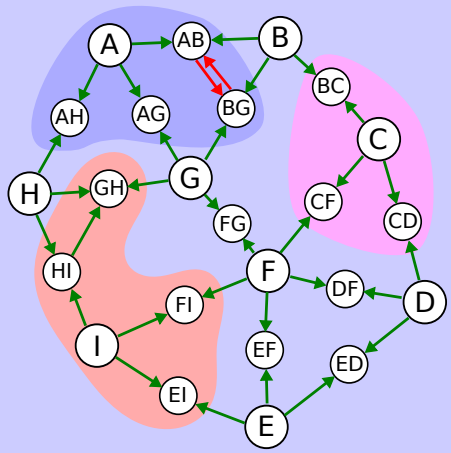


Prediction of Functional Importance

Minimal network state graph



Breadth first search from each node



Perturbation Impact Score

$$PIS_{(P,I,C)}(Q_{\downarrow}) := |BFS(Q_{\downarrow})|$$

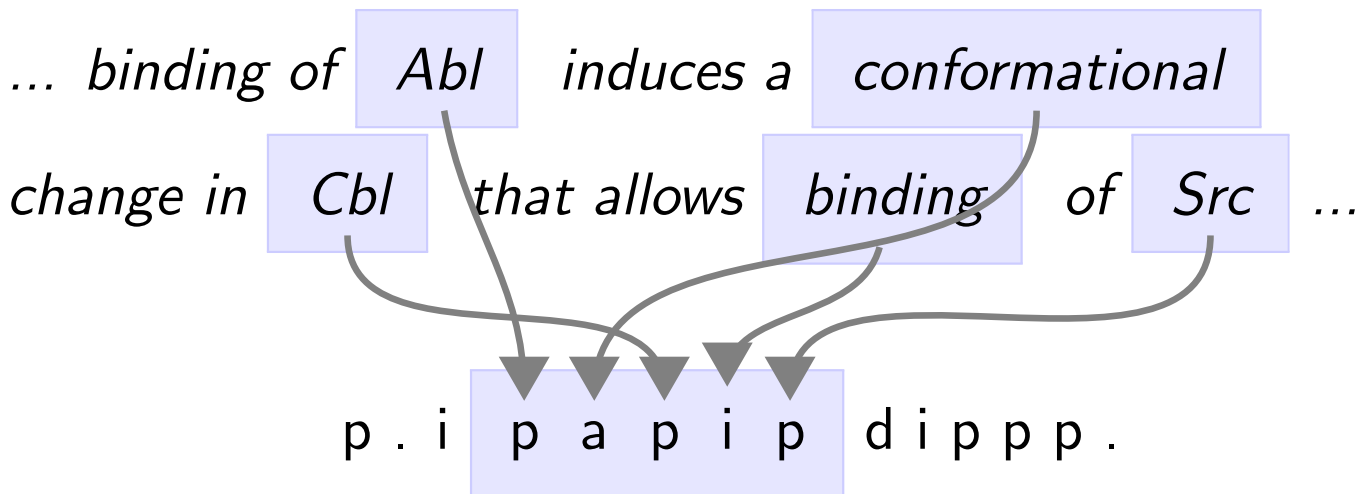
Harvesting Constraints

Text-Mining

- Observation: Interaction dependencies are reported as single sentence natural language statements in literature.
- Tokenize full-text papers into relevant words and search for simple regular expression patterns.

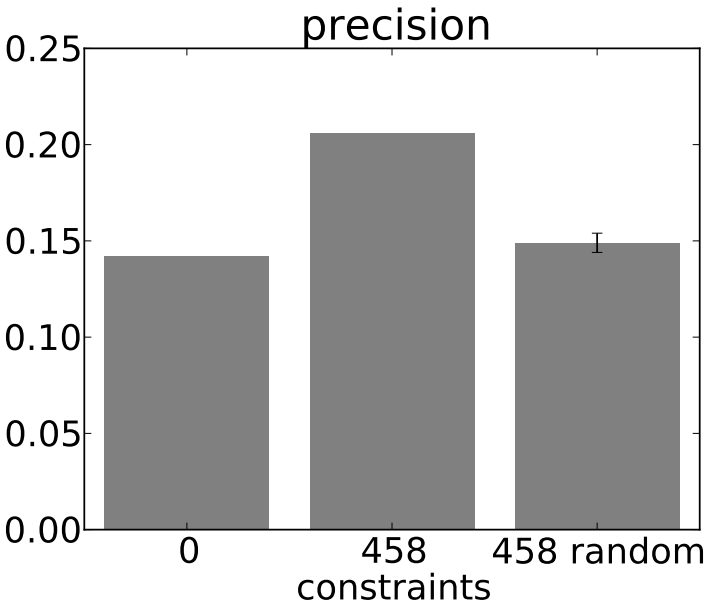
71 new interaction dependencies from 59 000 human adhesome related papers.

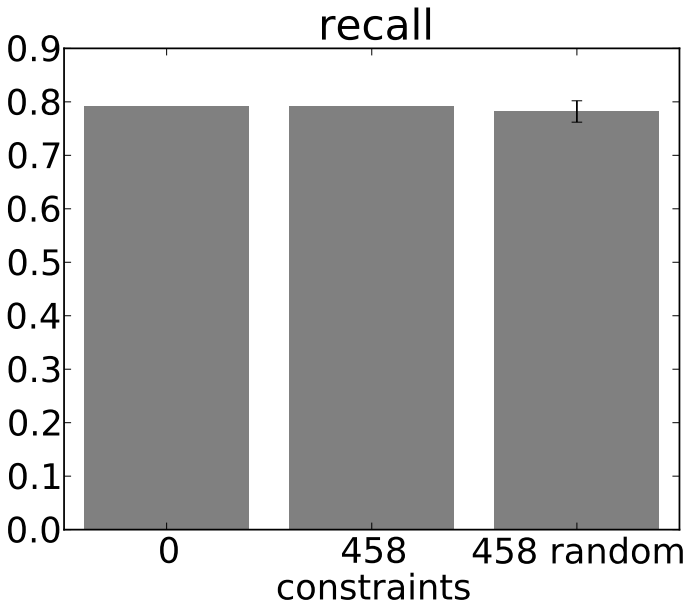
Köster, Zamir, Rahmann. 2012



Results for Complex Predicton

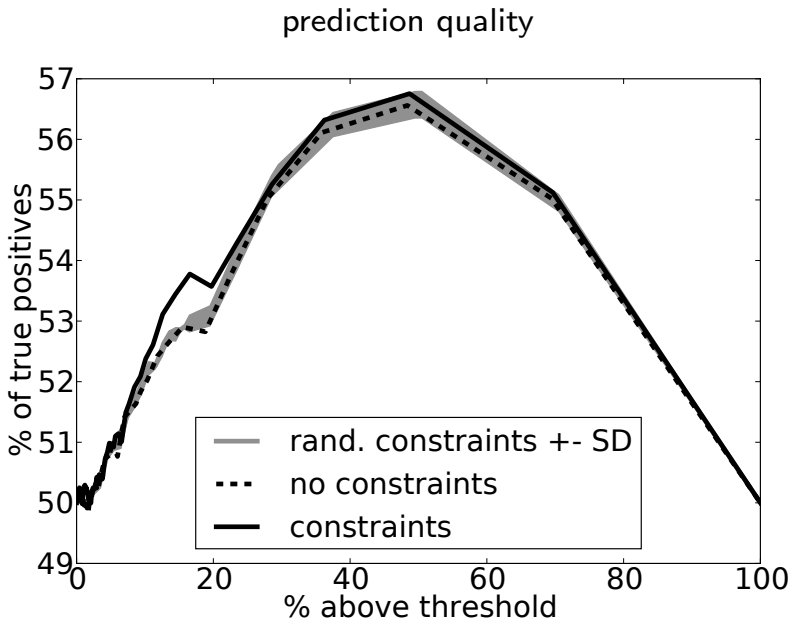
- Network: CYGD (4579 proteins, 12576 interactions)
- Constraints: Competition on binding sites *(Jung et al. 2010)*
- Complexes: CYGD (55 connected complexes)
- Network based complex prediction: LCMA *(Li et al. 2005)*





Results for Perturbation Impact Score

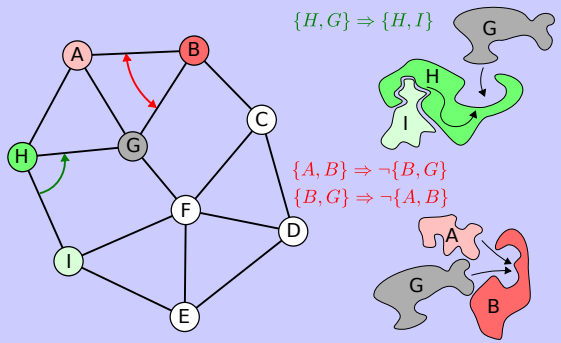
- Network: CYGD (4579 proteins, 12576 interactions)
- Constraints: Competition on binding sites (Jung et al. 2010)
- Perturbations classified as lethal/sick and viable: SGD



TP: lethal/sick and $PIS \geq t$, viable and $PIS < t$

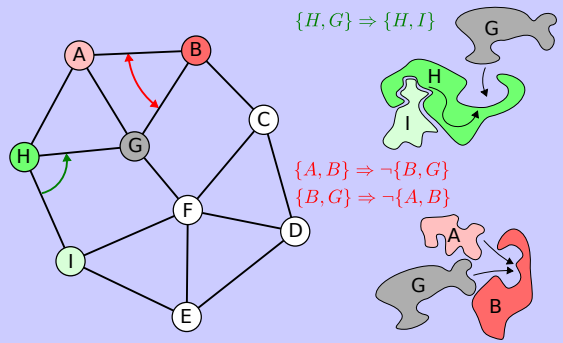
Conclusion

Protein Hypernetworks



Conclusion

Protein Hypernetworks



more precise protein complexes

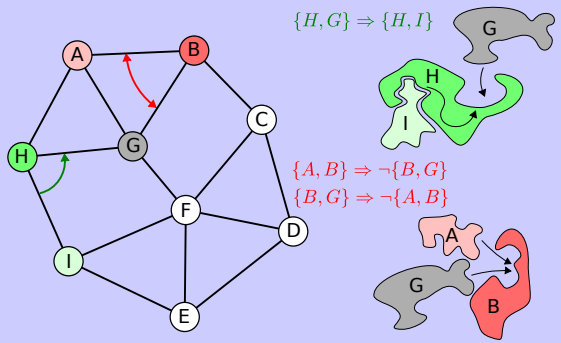
perturbation effects

Conclusion

text-mining

logic inference from real measurements

Protein Hypernetworks



more precise protein complexes

perturbation effects

Harvesting Constraints

Using the Quine-McCluskey-Algorithm

- Given a truth table with interactions in columns and simultaneous observations in rows.
- Infer logic relationships using the Quine-McCluskey-Algorithm.

AB	BC	observed
0	0	1
0	1	1
1	0	1
1	1	0

Inferred constraints:
$AB \Rightarrow \neg BC$

Harvesting Constraints

Using the Quine-McCluskey-Algorithm

- Given a truth table with interactions in columns and simultaneous observations in rows.
- Infer logic relationships using the Quine-McCluskey-Algorithm.

derive rows from

- simultaneous interaction measurements (e.g. future variants of FCS)
- combination of protein complex measurements (e.g. MS) with binary protein interactions

AB	BC	observed
0	0	1
0	1	1
1	0	1
1	1	0

Inferred constraints:
$AB \Rightarrow \neg BC$