Theorem Proving

Three important concepts for Logical Inference

Horn Clauses

* ### A disjunction in which at most one literal is positive

* ### Equivalent in form to an implication

* ### Example

$$\neg L_{1,1} \vee \neg Breeze \vee B_{1,1} \Longrightarrow (L_{1,1} \wedge Breeze) \Rightarrow B_{1,1}$$

$$P_{1,1} \vee \neg P_{2,1} \vee B_{1,2}$$

Inference Algorithms: Forward and Backward Chaining

Forward and Backward Chaining

Horn Form (restricted)
$$KB = \textbf{conjunction of Horn clauses}$$
Horn clause =
$\diamond$ proposition symbol; or
$\diamond$ (conjunction of symbols) $\Rightarrow$ symbol
E.g., $C \wedge (B \Rightarrow A) \wedge (C \wedge D \Rightarrow B)$

* Restrict sentences to be written a particular way

* Only need one inference rule in your search for entailment

Forward and Backward Chaining

Horn Form (restricted)

KB = **conjunction** of **Horn clauses**

Horn clause =

◇ proposition symbol; or

◇ (conjunction of symbols) $\Rightarrow$ symbol

E.g., $C \wedge (B \Rightarrow A) \wedge (C \wedge D \Rightarrow B)$
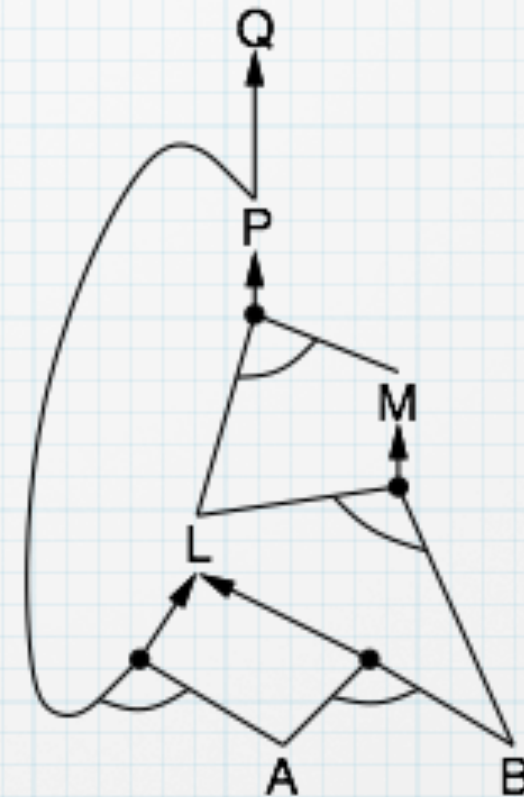
Modus Ponens (for Horn Form): complete for Horn KBs

$$\alpha_1, \ldots, \alpha_n, \qquad \alpha_1 \wedge \cdots \wedge \alpha_n \Rightarrow \beta$$
$$\beta$$

Can be used with forward chaining or backward chaining.
These algorithms are very natural and run in **linear** time

Forward Chaining

Idea: fire any rule whose premises are satisfied in the KB,
add its conclusion to the KB, until query is found

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
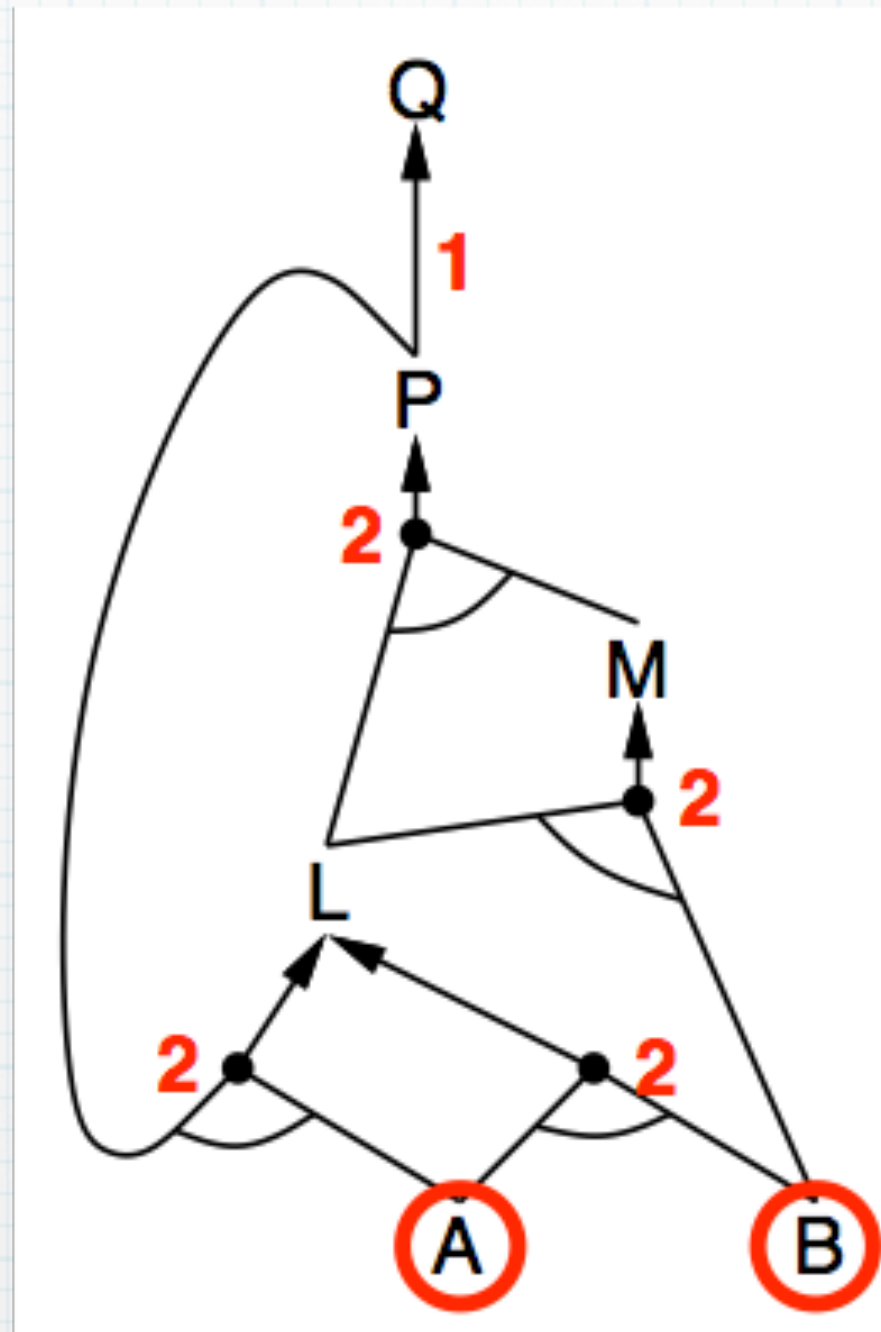$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

Forward Chaining

```
function PL-FC-ENTAILS?(KB, q) returns true or false
    inputs: KB, the knowledge base, a set of propositional Horn clauses
            q, the query, a proposition symbol
    local variables: count, a table, indexed by clause, initially the number of premises
                     inferred, a table, indexed by symbol, each entry initially false
                     agenda, a list of symbols, initially the symbols known in KB

    while agenda is not empty do
        p ← POP(agenda)
        unless inferred[p] do
            inferred[p] ← true
            for each Horn clause c in whose premise p appears do
                decrement count[c]
                if count[c] = 0 then do
                    if HEAD[c] = q then return true
                    PUSH(HEAD[c], agenda)
    return false
```

Forward chaining example

$$P \Rightarrow Q$$
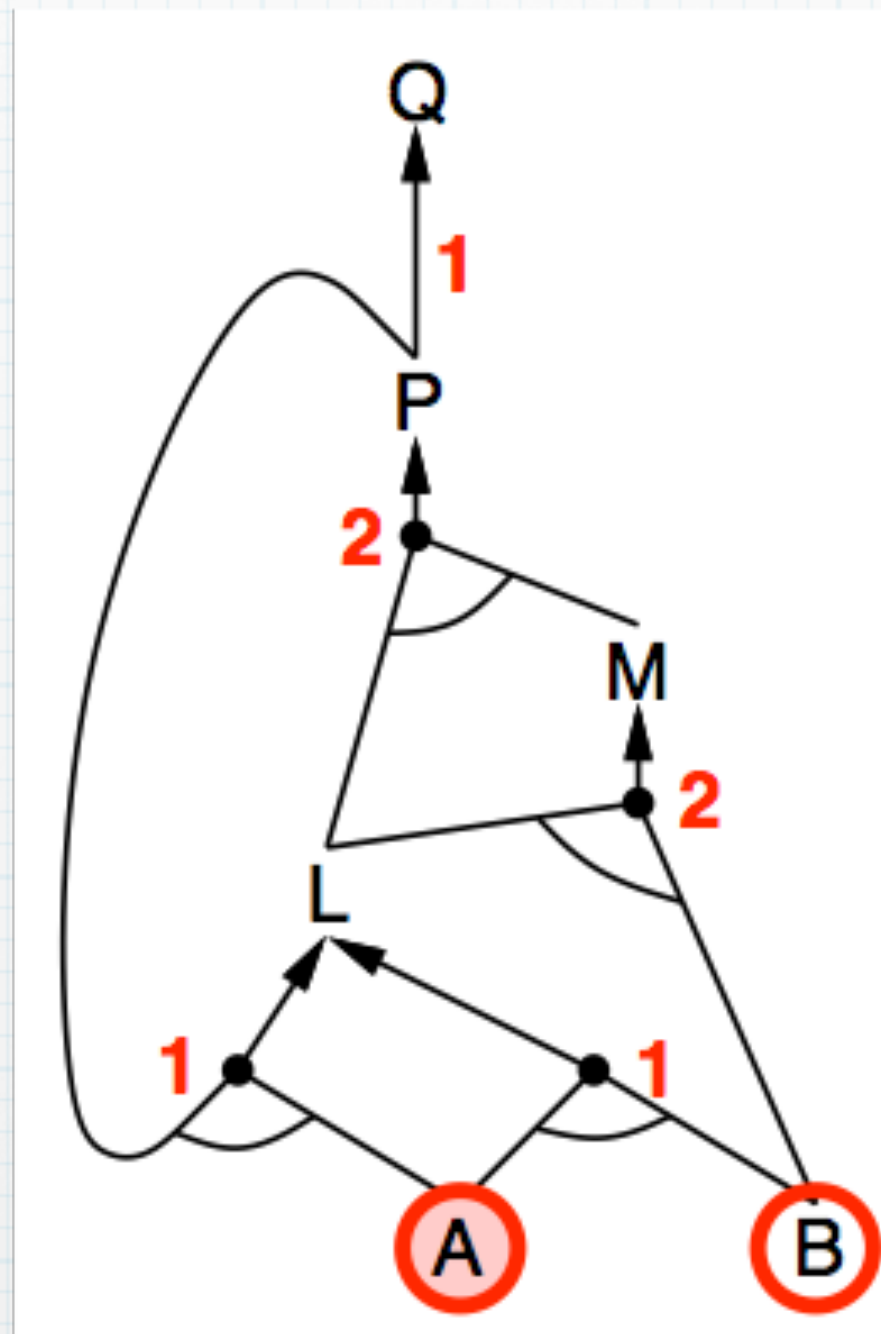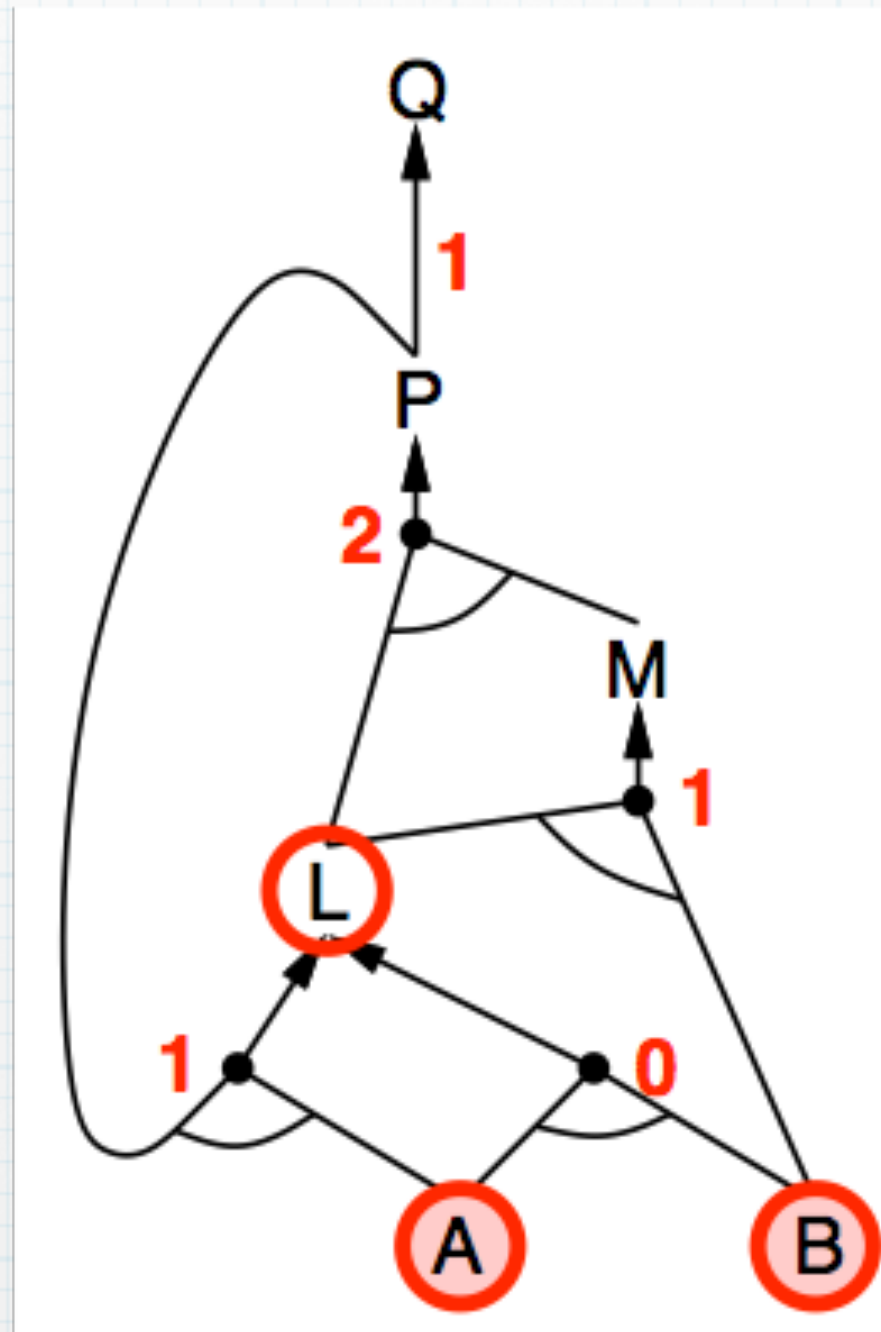$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

Forward chaining example

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
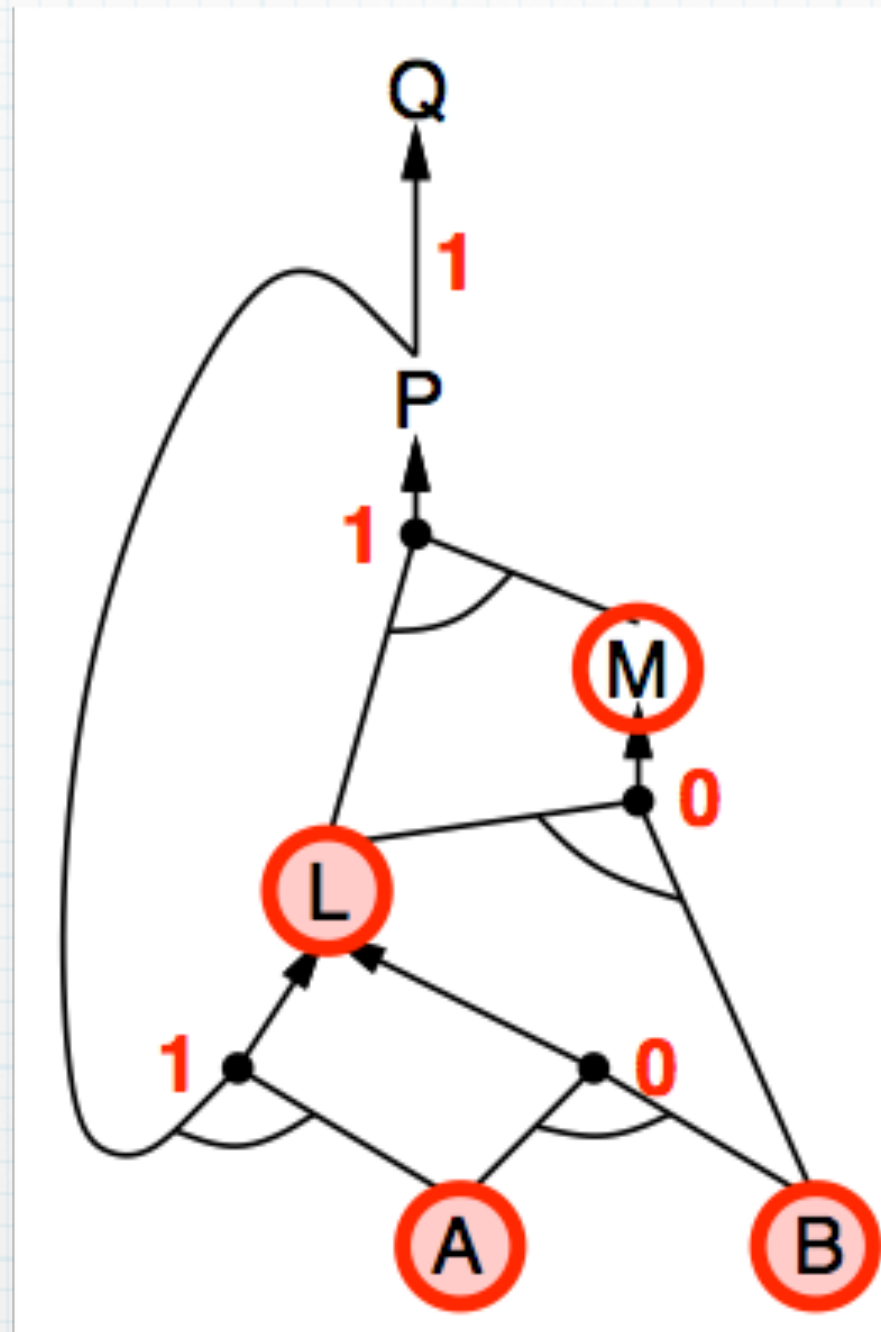$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

Forward chaining example

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
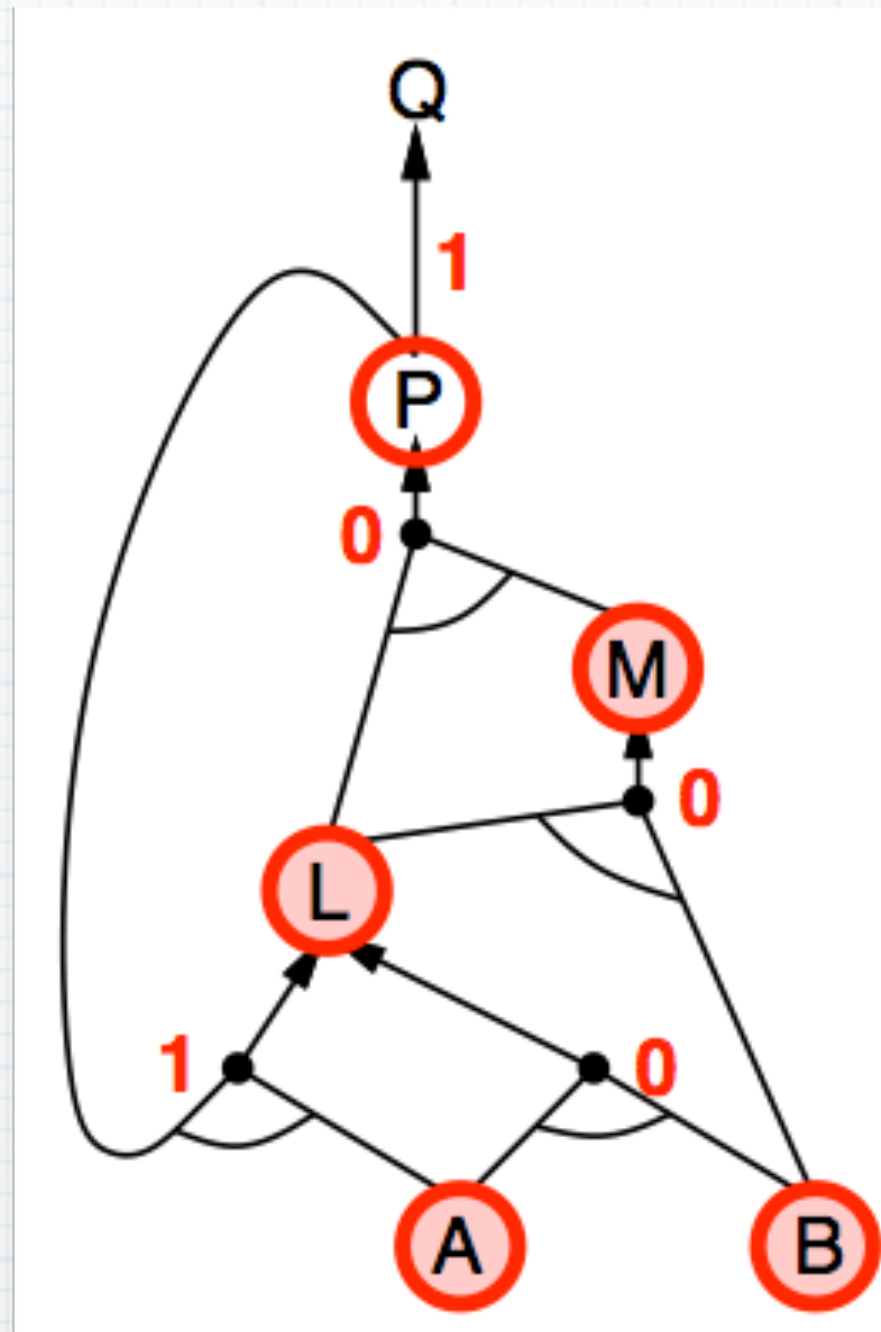$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$
___

L

Forward chaining example

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
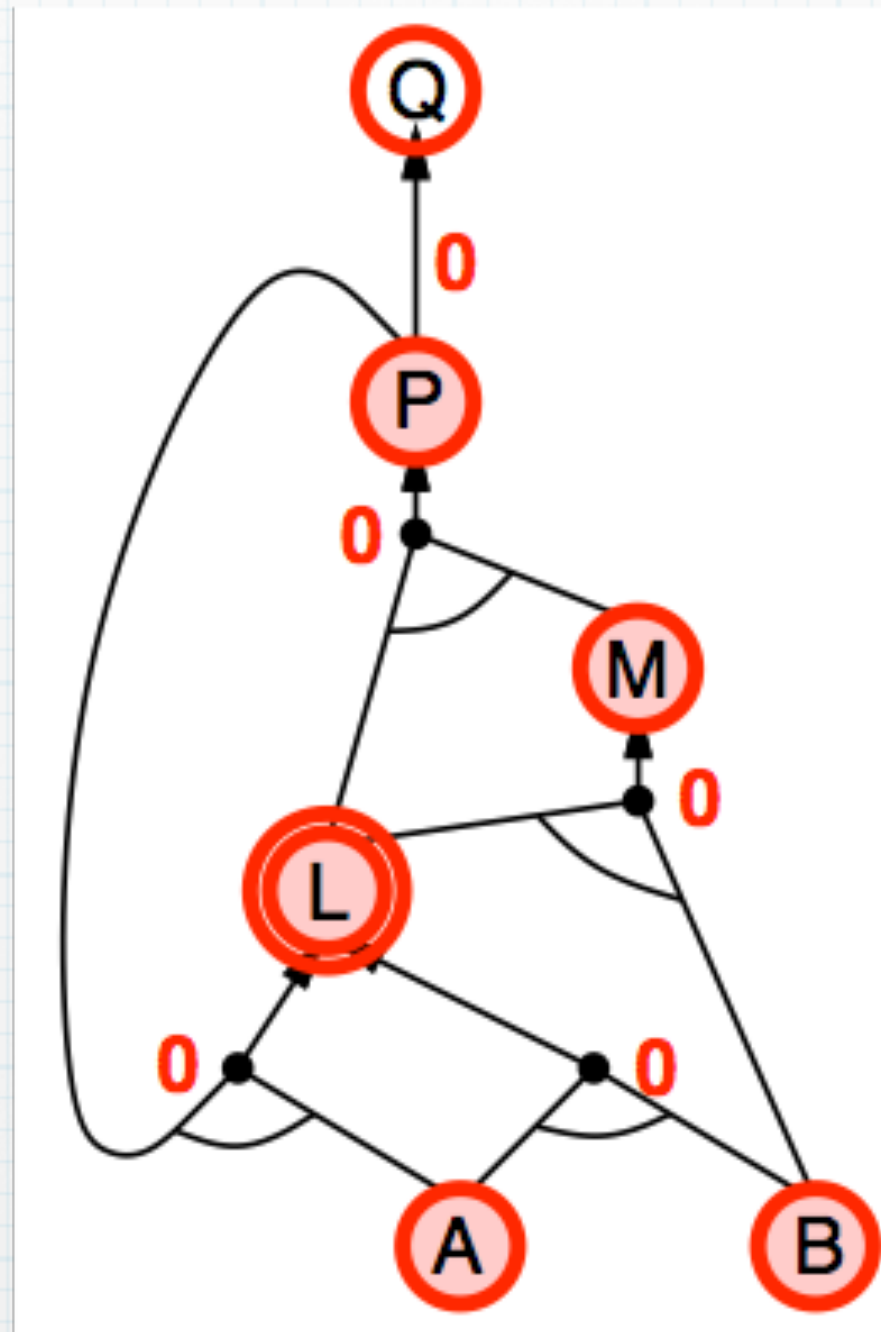$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

$$L$$

Forward chaining example

$P \Rightarrow Q$

$L \wedge M \Rightarrow P$

$B \wedge L \Rightarrow M$

$A \wedge P \Rightarrow L$

$A \wedge B \Rightarrow L$

A

B

L

M

Forward chaining example

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
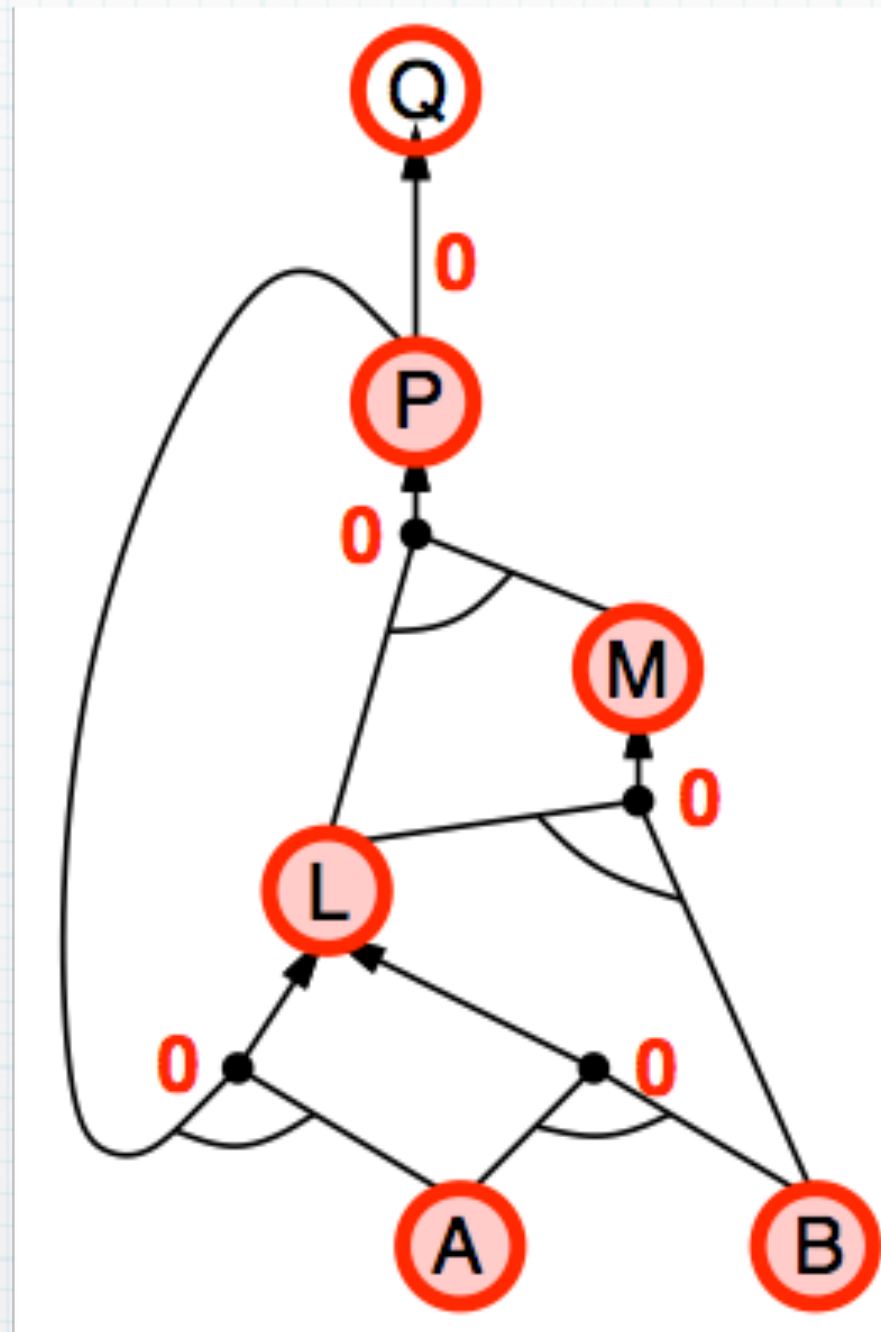$$A \wedge P \Rightarrow L$$
$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

L

M

P

Forward chaining example

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
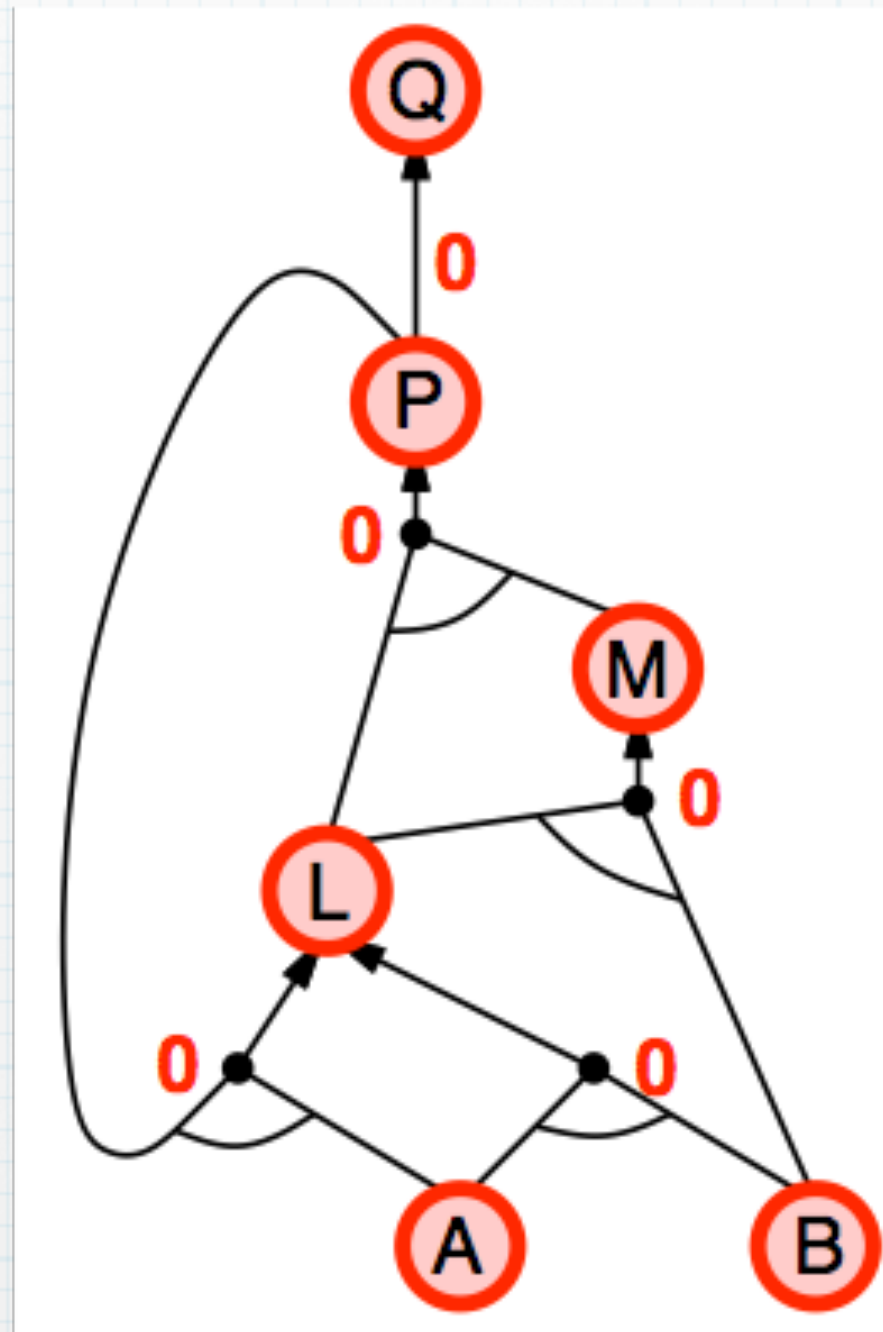$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

L

M

P

Forward chaining example

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

L
M
P
Q

Proof of Completeness

FC derives every atomic sentence that is entailed by KB

1. FC reaches a fixed point where no new atomic sentences are derived

2. Consider the final state as a model m, assigning true/false to symbols

3. Every clause in the original KB is true in m
   Proof: Suppose a clause $a_1 \wedge \ldots \wedge a_k \Rightarrow b$ is false in m
   Then $a_1 \wedge \ldots \wedge a_k$ is true in m and b is false in m
   Therefore the algorithm has not reached a fixed point!

4. Hence m is a model of KB

5. If $KB \models q$, q is true in **every** model of KB, including m

General idea: construct any model of KB by sound inference, check α

Backward Chaining

Idea: work backwards from the query q:
    to prove q by BC,
        check if q is known already, or
        prove by BC all premises of some rule concluding q

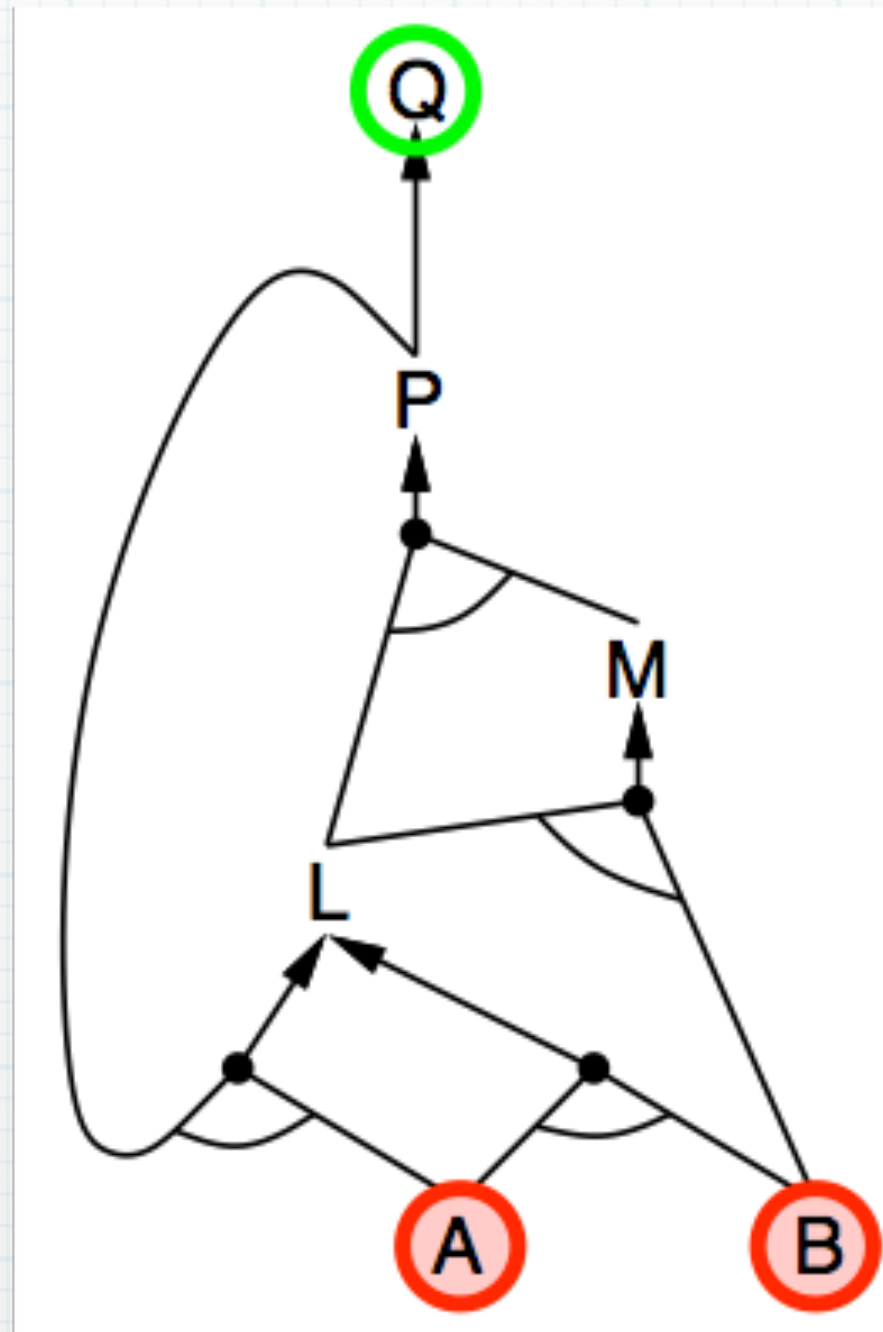Avoid loops: check if new subgoal is already on the goal stack
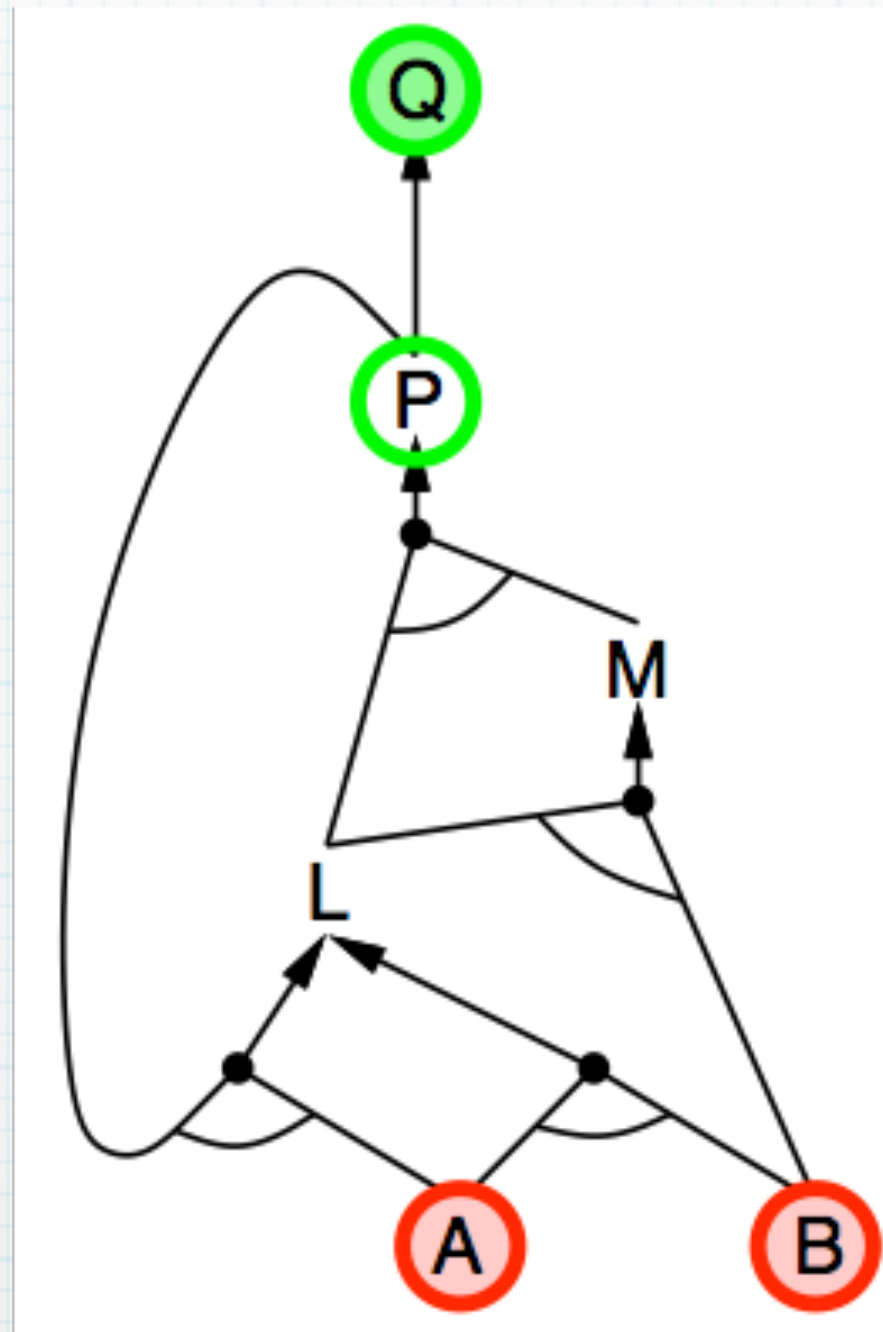
Avoid repeated work: check if new subgoal
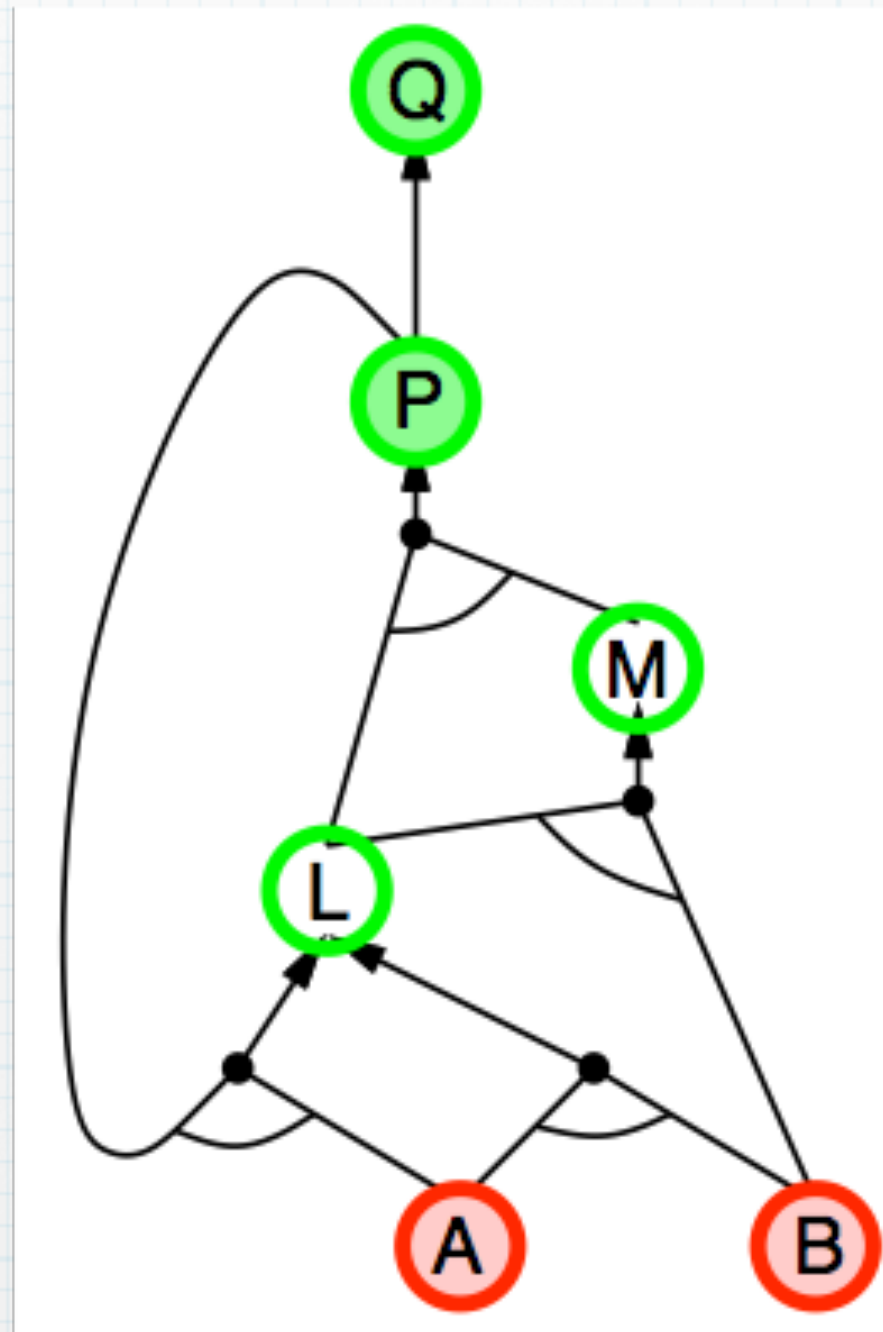   1) has already been proved true, or
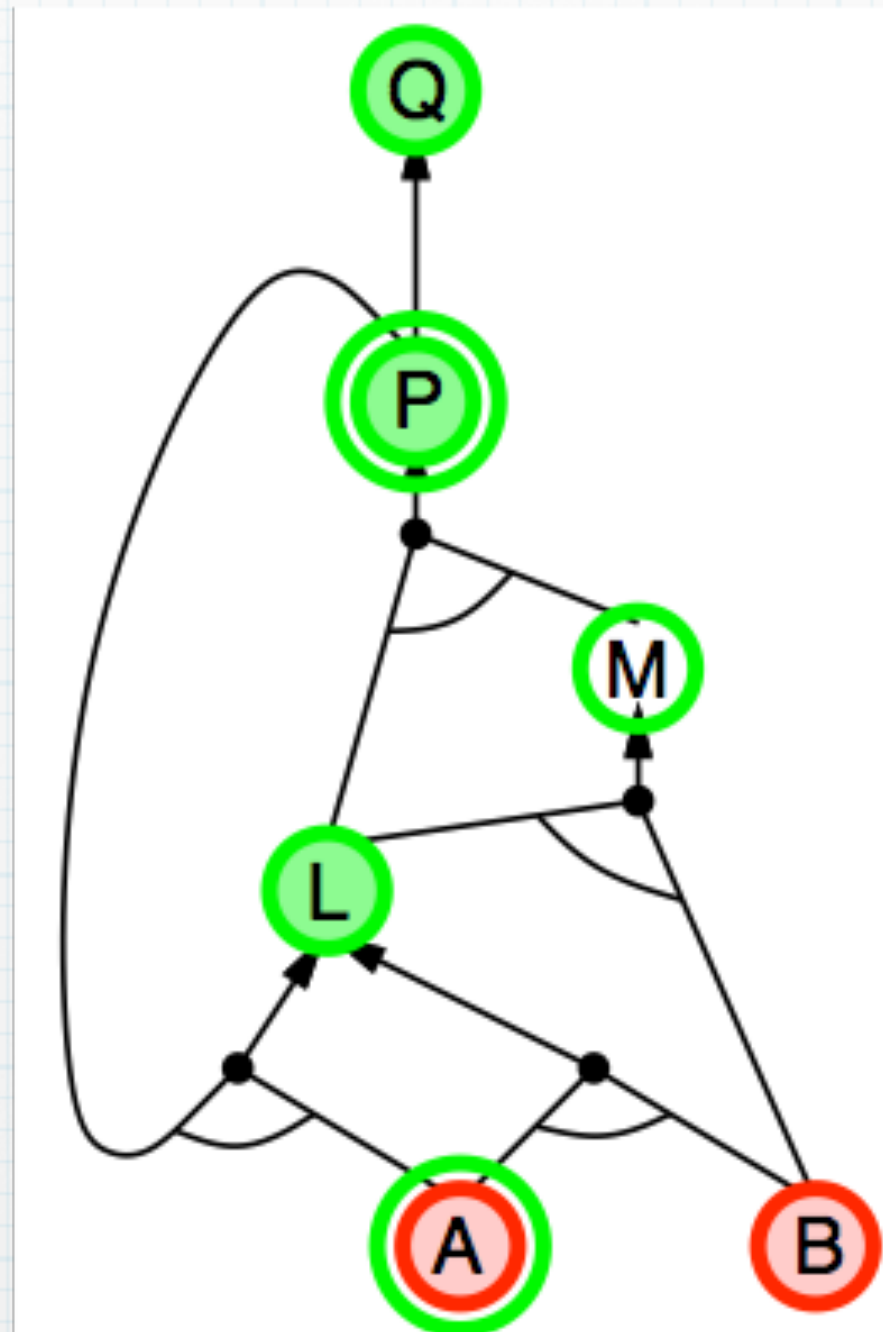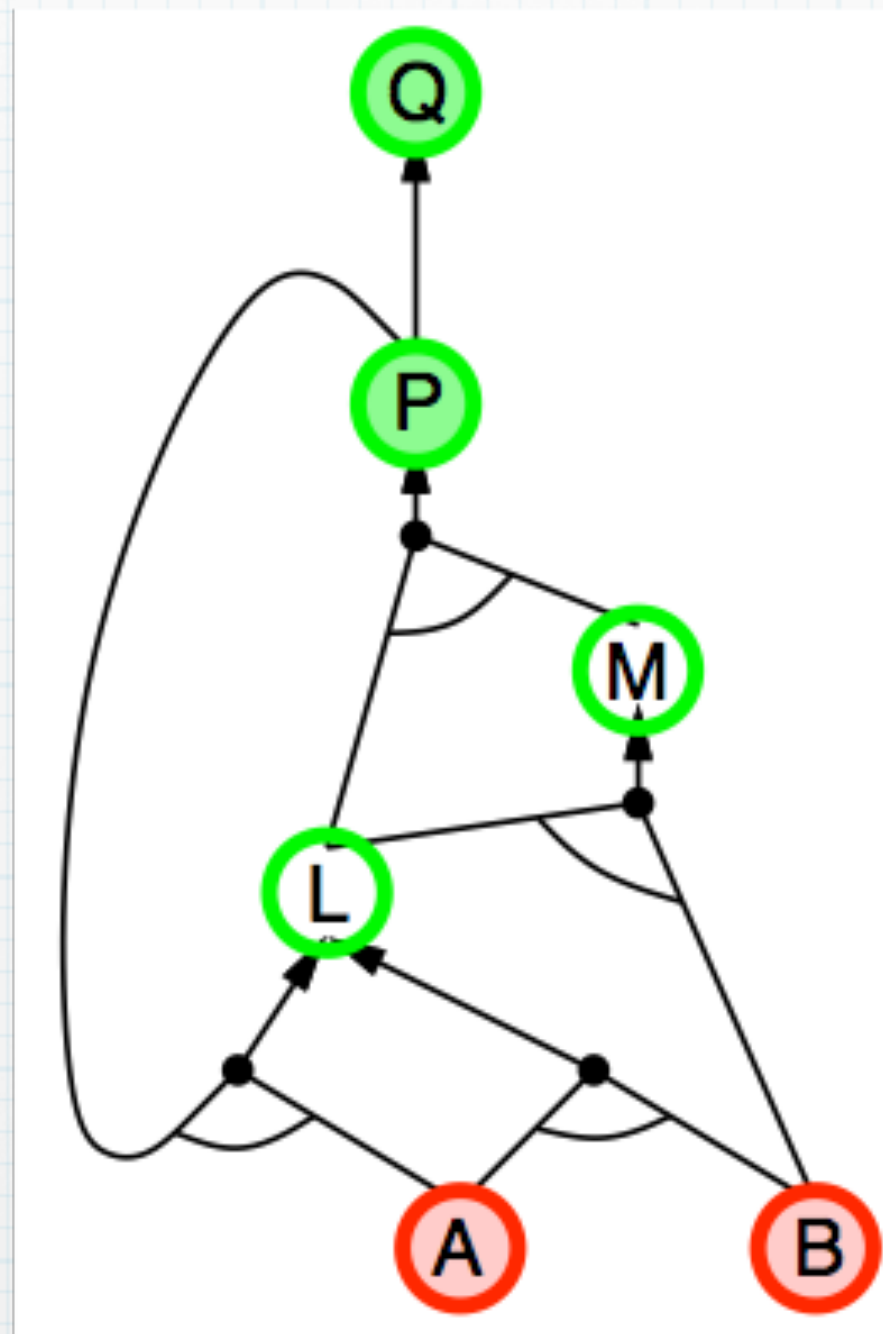   2) has already failed

Backward chaining example

Backward chaining example

Backward chaining example

Backward chaining example

Backward chaining example

Backward chaining example
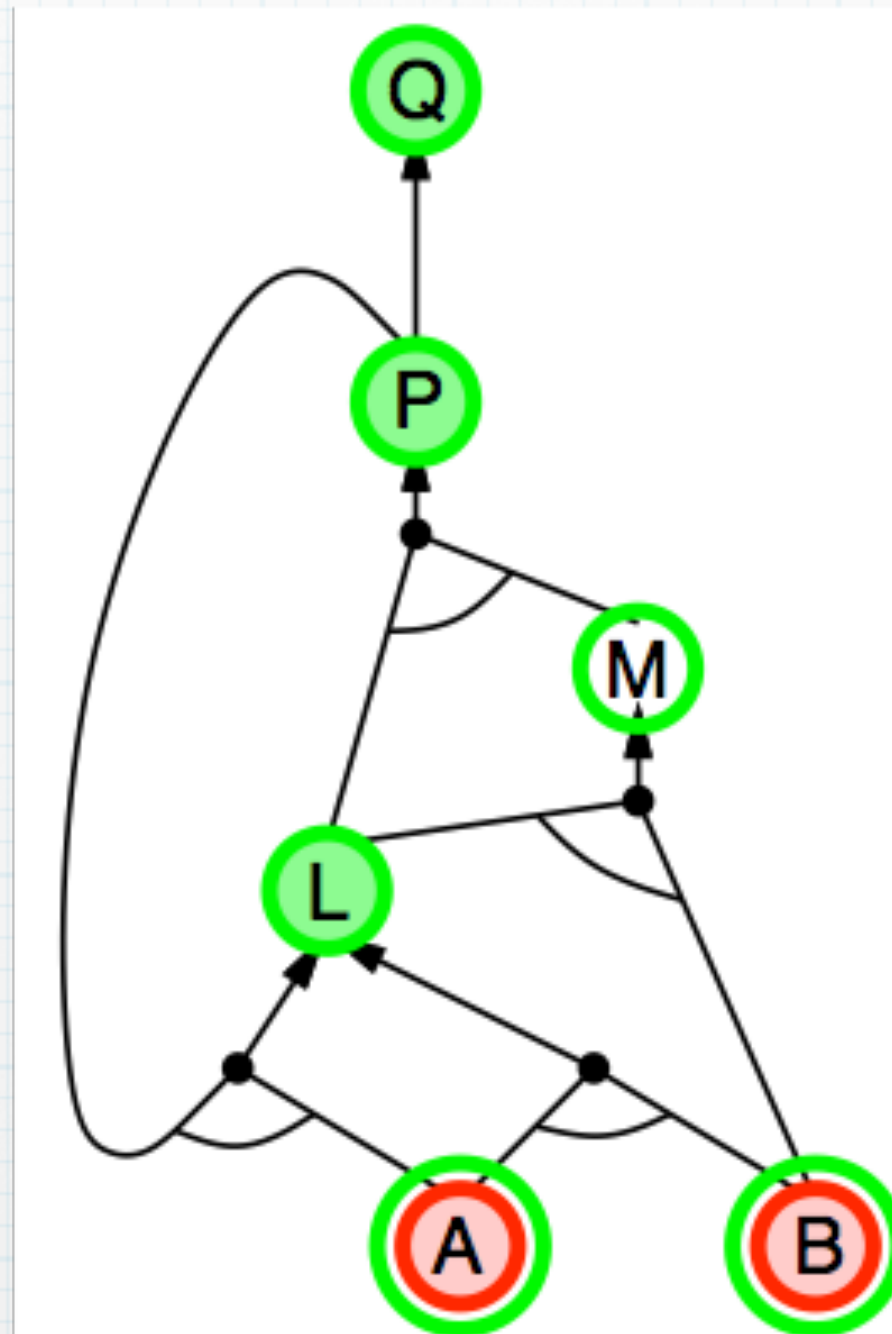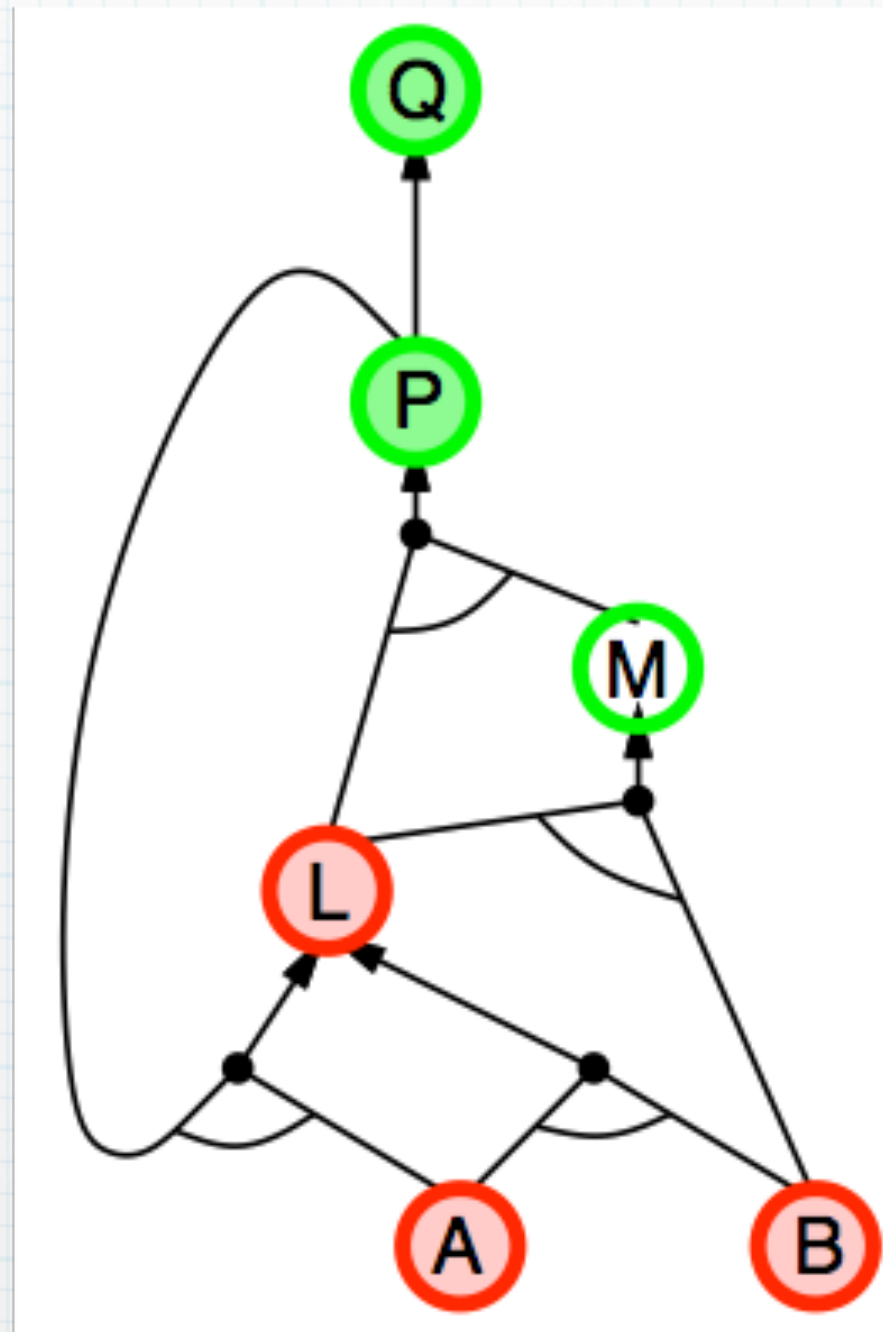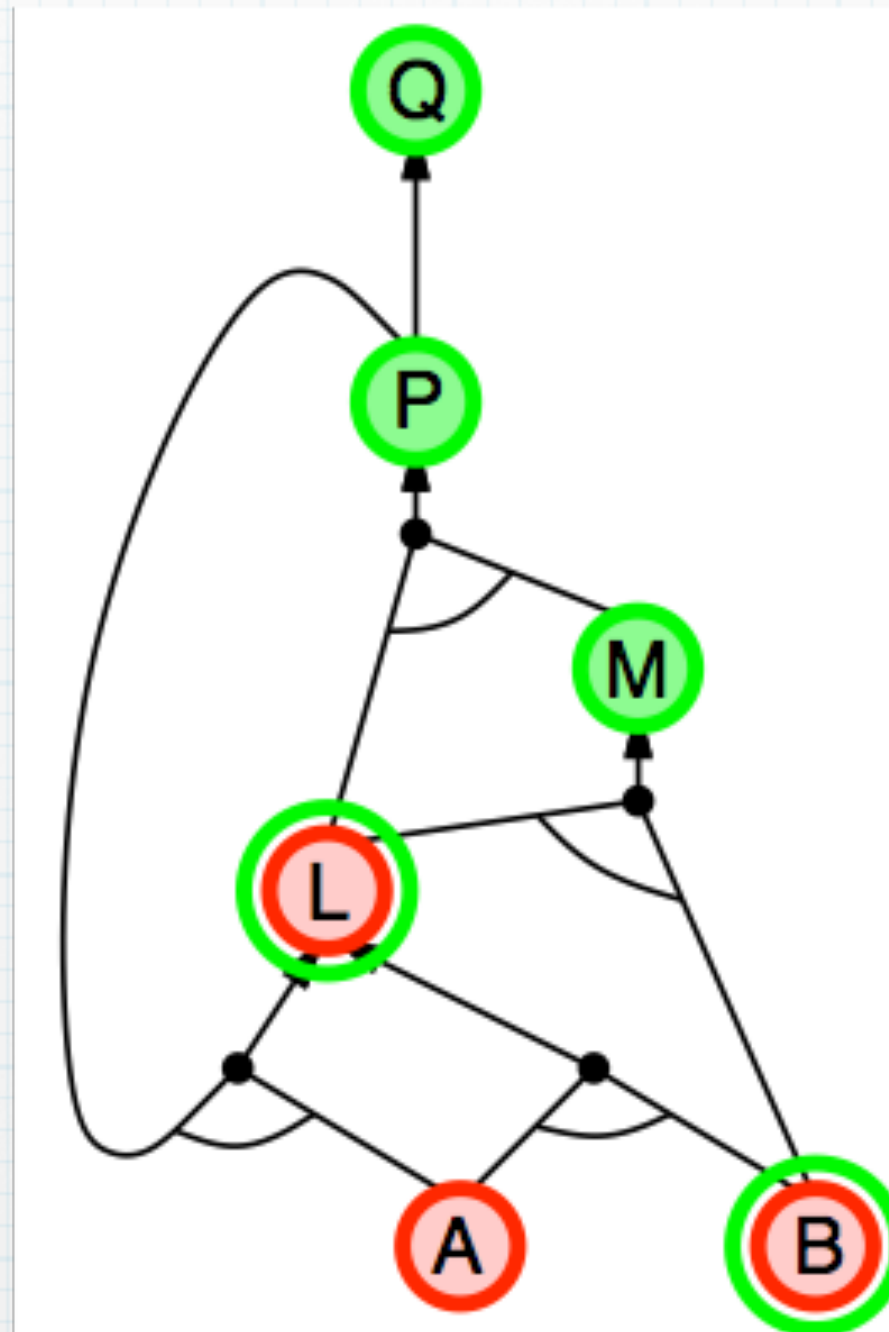
Backward chaining example

Backward chaining example

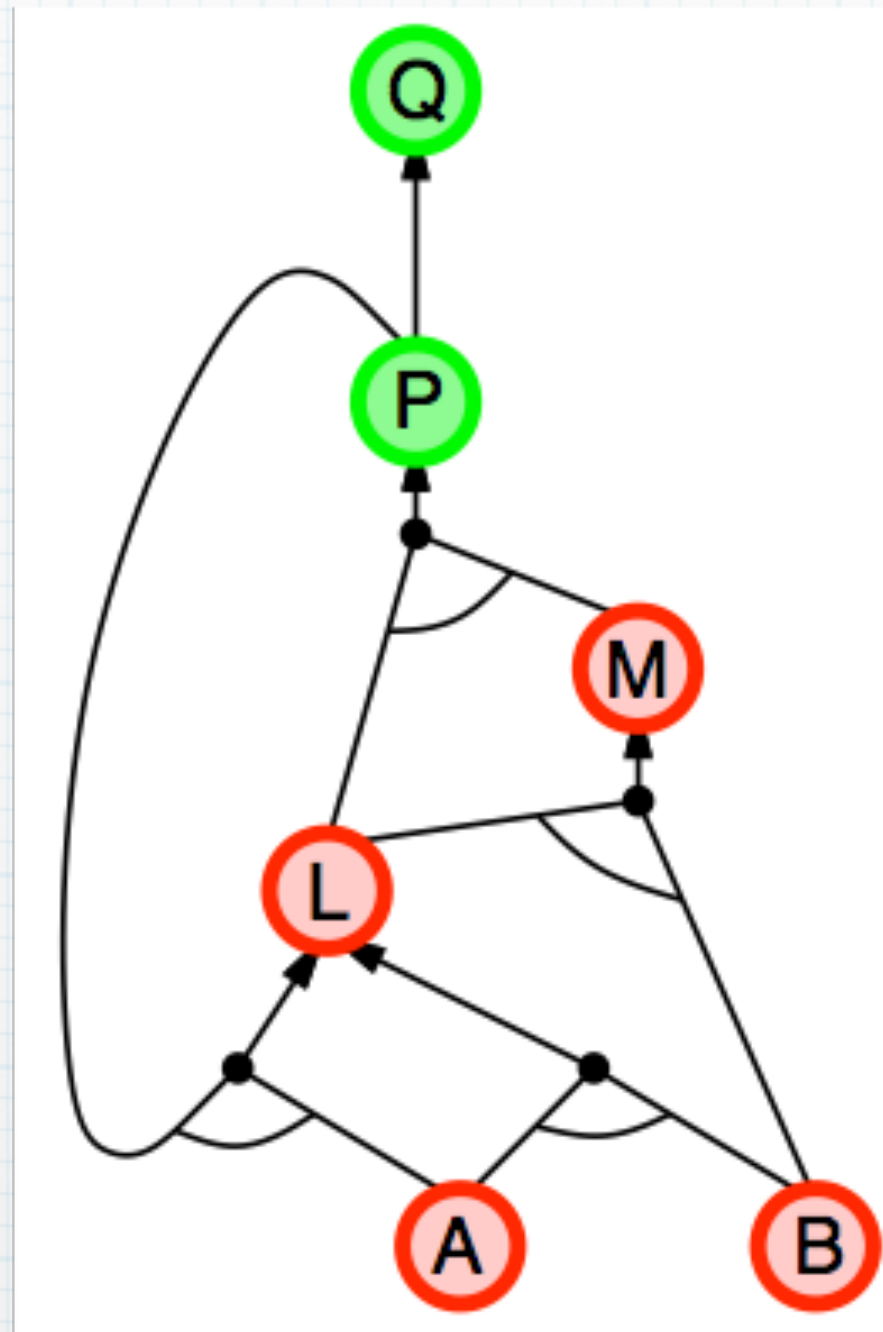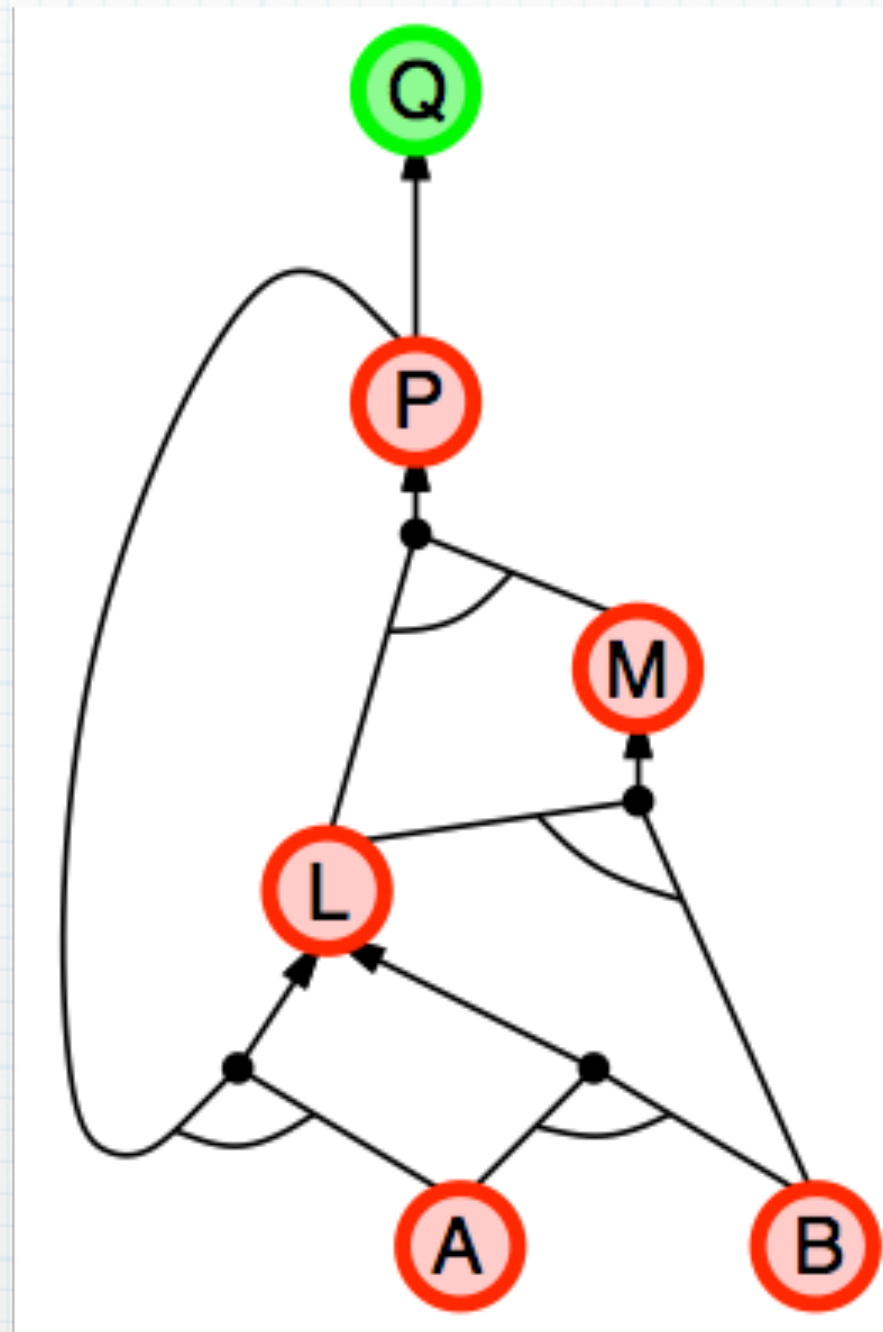# Backward chaining example

Backward chaining example

Backward chaining example
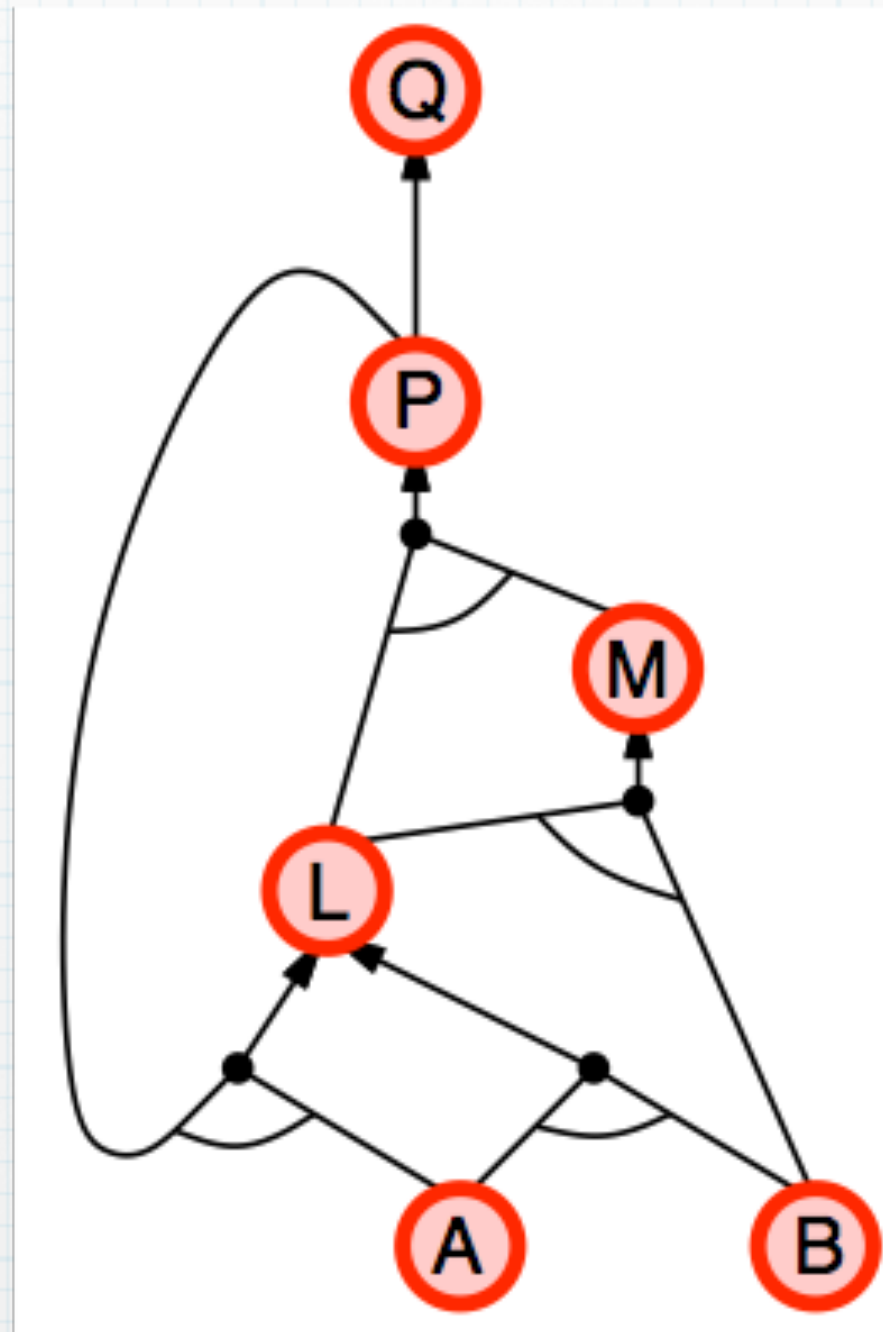
Forward vs. Backward Chaining

FC is data-driven, cf. automatic, unconscious processing,
        e.g., object recognition, routine decisions

May do lots of work that is irrelevant to the goal

BC is goal-driven, appropriate for problem-solving,
        e.g., Where are my keys? How do I get into a PhD program?

Complexity of BC can be **much less** than linear in size of KB

Inference Algorithms: Resolution

Resolution

Conjunctive Normal Form (CNF—universal)

$$\underbrace{\text{conjunction} \text{ of } \overbrace{\text{disjunctions} \text{ of } \text{literals}}^{\text{clauses}}}$$

E.g., $(A \vee \neg B) \wedge (B \vee \neg C \vee \neg D)$

Resolution

Conjunctive Normal Form (CNF—universal)

$$\text{conjunction of } \underbrace{\text{disjunctions of literals}}_{\text{clauses}}$$

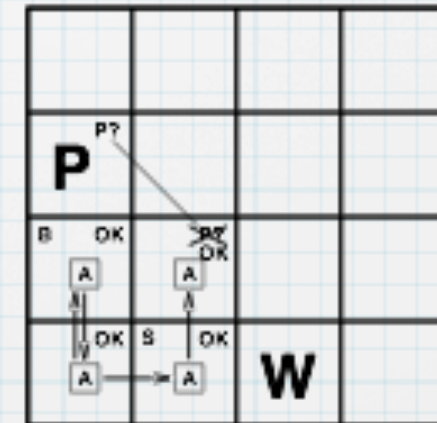E.g., $(A \vee \neg B) \wedge (B \vee \neg C \vee \neg D)$

Resolution inference rule (for CNF): complete for propositional logic

$$\frac{\ell_1 \vee \cdots \vee \ell_k, \qquad m_1 \vee \cdots \vee m_n}{\ell_1 \vee \cdots \vee \ell_{i-1} \vee \ell_{i+1} \vee \cdots \vee \ell_k \vee m_1 \vee \cdots \vee m_{j-1} \vee m_{j+1} \vee \cdots \vee m_n}$$

where ℓ_i and m_j are complementary literals. E.g.,

$$\frac{P_{1,3} \vee P_{2,2}, \qquad \neg P_{2,2}}{P_{1,3}}$$

Resolution is sound and complete for propositional logic

Resolution Algorithm

Proof by contradiction, i.e., show $KB \wedge \neg \alpha$ unsatisfiable

Resolution Algorithm

Proof by contradiction, i.e., show $KB \wedge \neg\alpha$ unsatisfiable

---

function PL-RESOLUTION(KB, α) **returns** *true* or *false*
    inputs: KB, the knowledge base, a sentence in propositional logic
           α, the query, a sentence in propositional logic

   clauses ← the set of clauses in the CNF representation of $KB \wedge \neg\alpha$
   new ← { }
   loop do
      for each C_i, C_j **in** *clauses* **do**
         resolvents ← PL-RESOLVE(C_i, C_j)
         if *resolvents* contains the empty clause **then return** *true*
         new ← *new* ∪ *resolvents*
      if *new* ⊆ *clauses* **then return** *false*
   clauses ← *clauses* ∪ *new*

Resolution Example

$$KB = (B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})) \wedge \neg B_{1,1} \quad \alpha = \neg P_{1,2}$$

$\neg P_{2,1} \vee B_{1,1}$	$\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}$	$\neg P_{1,2} \vee B_{1,1}$	$\neg B_{1,1}$	$P_{1,2}$

Resolution Example

$$KB = (B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})) \wedge \neg B_{1,1} \quad \alpha = \neg P_{1,2}$$

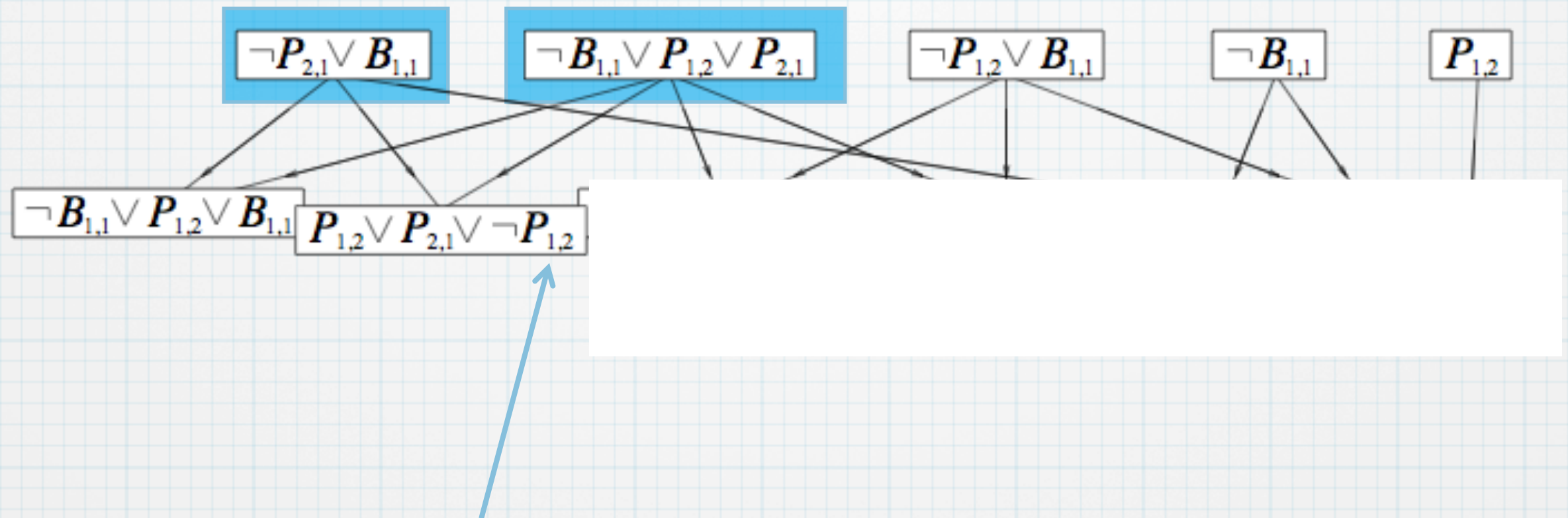| $\neg P_{2,1} \vee B_{1,1}$ | $\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}$ | $\neg P_{1,2} \vee B_{1,1}$ | $\neg B_{1,1}$ | $P_{1,2}$ |

$\neg B_{1,1} \vee P_{1,2} \vee B_{1,1}$ $P_{1,2} \vee P_{2,1} \vee \neg P_{1,2}$

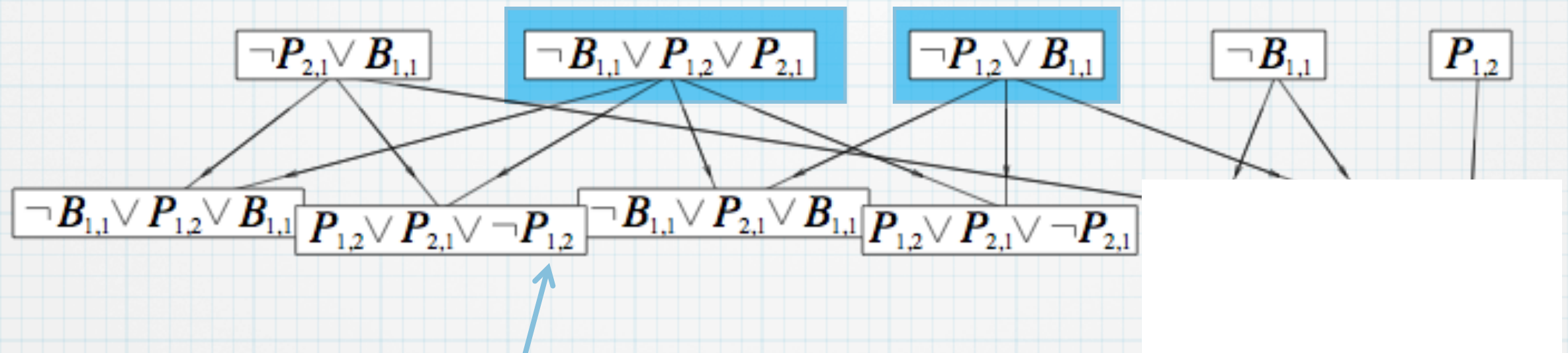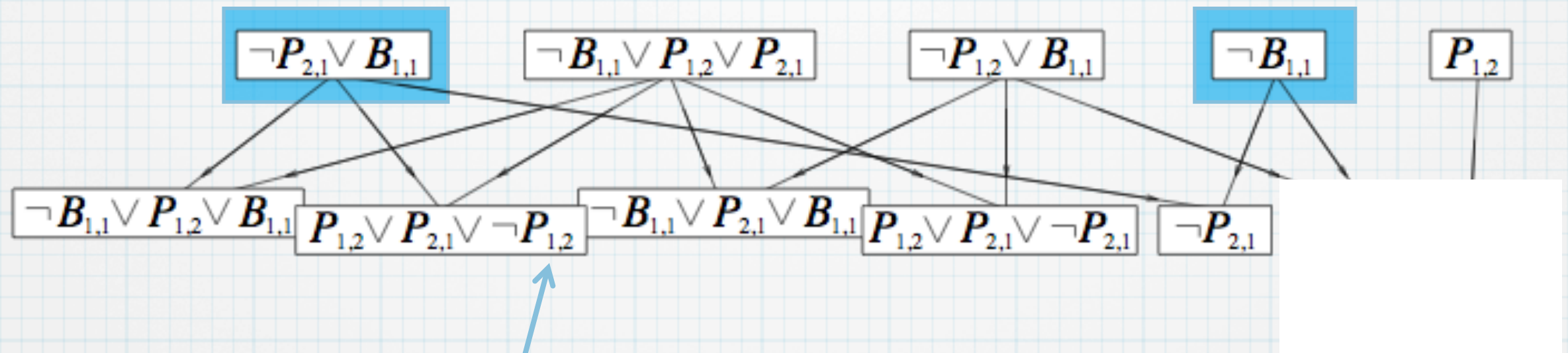**Note: Typo!! This should be** $\neg P \downarrow 2,1$

Resolution Example
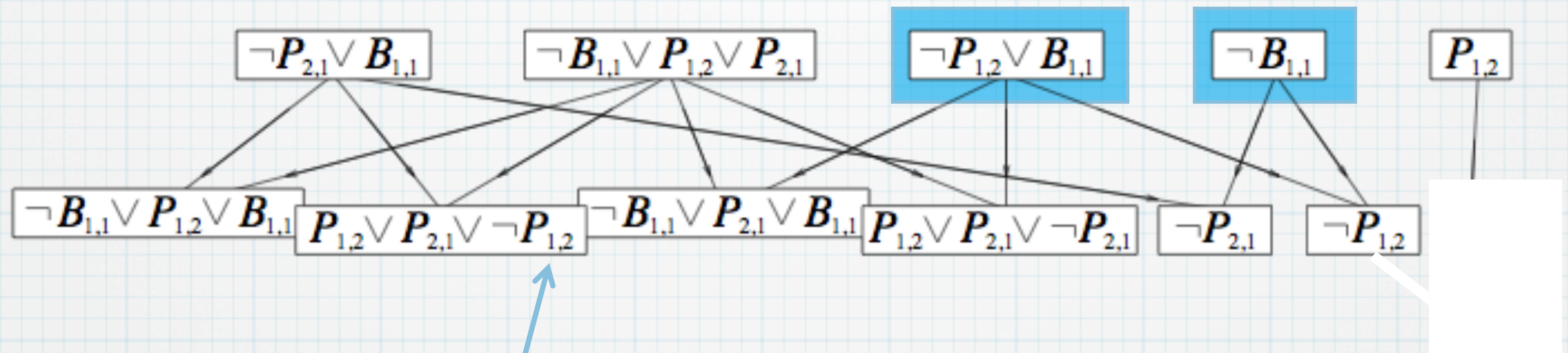
$$KB = (B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})) \wedge \neg B_{1,1} \quad \alpha = \neg P_{1,2}$$

Note: Typo!! This should be $\neg P{\downarrow}2,1$

Resolution Example

$$KB = (B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})) \wedge \neg B_{1,1} \quad \alpha = \neg P_{1,2}$$

Note: Typo!! This should be $\neg P{\downarrow}2,1$

Resolution Example
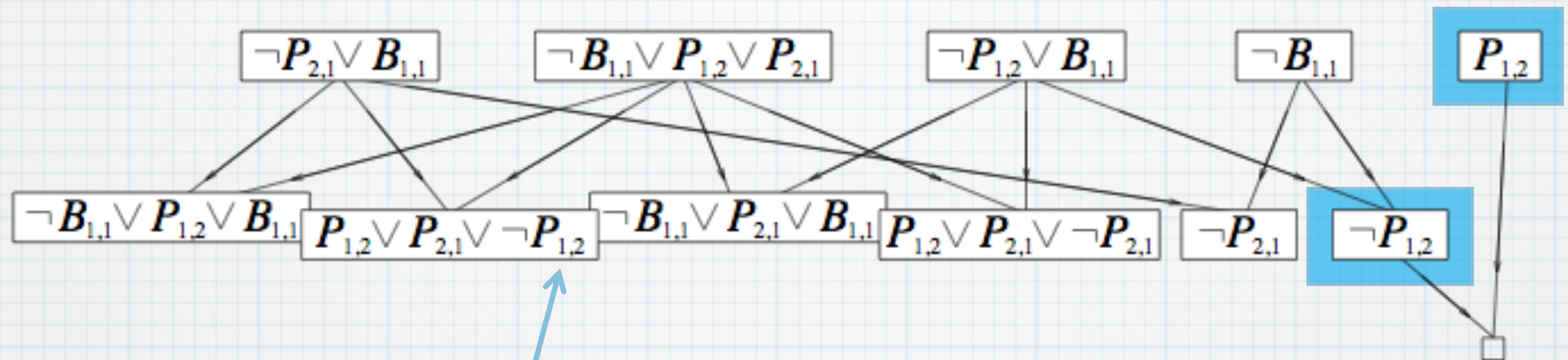
$$KB = (B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})) \wedge \neg B_{1,1} \quad \alpha = \neg P_{1,2}$$

Note: Typo!! This should be $\neg P{\downarrow}2,1$

Resolution Example

$$KB = (B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})) \land \neg B_{1,1} \quad \alpha = \neg P_{1,2}$$

Note: Typo!! This should be $\neg P\!\downarrow\!2,1$

Contradiction, therefore our sentence is entailed.