Part Seven: Causal Reasoning

- For a rational agent to be able to construct plans for making the environment more to its liking, it must be able to reason causally.
- In particular, it must be able to reason its way through the frame problem.

The Frame Problem

- Reasoning about what will change if an action is performed or some other change occurs often presupposes knowing what will not change.
  - Suppose I want the light to be on in the room.
  - I know that if I am at the location of the switch and I throw it, the light will come on.
  - The location of the switch is by the door.
  - I can go to that location by walking there.
  - So I plan to walk to that location and throw the switch.
  - This presupposes that the switch will still be there when I get to that location.

The Frame Problem

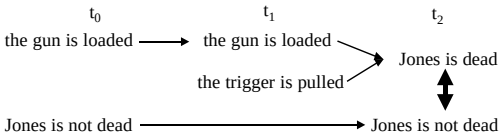
- Early attempts to model reasoning about change tried to do so deductively by adopting a large number of "frame axioms", which were axioms to the effect that if something occurs then something else will not change.
  - For instance, in a blocks world one of the frame axioms might be "If a block is moved, its color will not change".
- It soon became apparent that complicated situations required more frame axioms than axioms about change, and most of the system resources were being occupied by proofs that various properties did not change.
- What became known as the ***Frame Problem*** is the problem of reorganizing reasoning about change so that reasoning about non-change can be done efficiently (McCarthy and Hayes (1969).)

Reasoning Defeasibly about Non-Change

- Several authors (Sandewall (1972), McDermott (1982), McCarthy (1986)) proposed reasoning about change defeasibly and adopting some sort of defeasible inference scheme to the effect that it is reasonable to believe that something doesn't change unless you are forced to conclude otherwise.
- To make the idea work, one needs both a precise framework for defeasible reasoning and a precise formulation of the requisite defeasible inference schemes.
- The principle of TEMPORAL-PROJECTION and the OSCAR defeasible reasoner can be regarded as providing such a precise formulation.

The Frame Problem Resurrected

TEMPORAL-PROJECTION turns out to be only part of the solution, as was first shown by Hanks and McDermott (1986). ***The Yale Shooting Problem***



There is a kind of consensus that the solution to this problem lies in performing the temporal-projections in temporal order.

chronological minimalization — changes are minimized in chronological order

Chronological Minimalization

- Attempts to formalize chronological minimalization have met with mixed success, largely, I think, because they were based upon inadequate theories of defeasible reasoning.
- In addition, Kautz (1986) proposed a troublesome counterexample which seems to show that there is something wrong with the fundamental idea underlying chronological minimalization.
  - Suppose I leave my car in a parking lot at time t_0. I return at time t_3 to find it missing. Suppose I know somehow that it was stolen either at time t_1 or time t_2, where $t_0 < t_1 < t_2 < t_3$. Intuitively, there should be no reason to favor one of these times over the other as the time the car was stolen.
  - However, chronological minimalization would have us use temporal-projection first at t_1 to conclude that the car was still in the lot, and then because the car was stolen at either t_1 or t_2, we can conclude that the car was stolen at t_2.
  - This seems unreasonable.

Chronological Minimalization

- The difference between the cases in which chronological minimalization gives the intuitively correct answer and the cases in which it does not seems to be that in the former there is a set of temporal-projections that are rendered inconsistent by a causal connection between the propositions being projected.
- In the example of the stolen car, there is a set of temporal-projections not all of which can be correct, but the inconsistency does not result simply from a causal connection.
- The shooting case is causal, but the stolen car case is not.

Chronological Minimalization and Causal Undercutting

- When reasoning about such a causal system, part of the force of describing it as causal must be that the defeasible presumption against the effect occurring is somehow removed.
  - Thus, although we normally expect Jones to remain alive, we do not expect this any longer when he is shot.
- To remove a defeasible presumption is to defeat it.
- This suggests that there is some kind of general "causal" defeater for TEMPORAL PROJECTION:

For every $\varepsilon \geq 0$ and $\delta > 0$, "*A*&*P*-at-(t+ε) & (*A*&*P* causes *Q*)" is an undercutting defeater for the defeasible inference from ~*Q*-at-t to ~*Q*-at-(t+ε+δ) by **TEMPORAL-PROJECTION.**

Causation and Nomic Generalizations

- Causal undercutting cannot be correctly formulated in terms of "causes".
  - Causal overdetermination precludes the attribution of causes, but should not effect our reasoning about what will happen.
- "Causal laws" are formulated in terms of ***nomic generalizations***:

  (P ⇒ Q) means "Any physically possible P would be a Q"

  "(x is an electron ⇒ x is negatively charged)" means "(Any physically possible electron would be negatively charged."

  '⇒' is a variable-binding operator.

The logic of nomic generalizations is discussed at length in my ***Nomic Probability and the Foundations of Induction*** (Oxford, 1990).

Causal Undercutting

- Let us define "*A* when *P* is causally sufficient for *Q* after an interval ε" to mean

  $(\forall t)\{(A\text{-at-}t \ \& \ P\text{-at-}t) \ +\!> \ (\exists\delta)Q\text{-throughout-}(t+\varepsilon\ ,\ t+\varepsilon+\delta]\}$.

CAUSAL-UNDERCUTTER

Where $t_0 \leq t_1$ and $(t_1+\varepsilon) < t$, "*A*-at-t_1 & *Q*-at-t_1 & (*A* when *Q* is causally sufficient for ~*P* after an interval ε)" is a defeasible undercutting defeater for the inference from *P*-at-t_0 to P-throughout-(t^*,t) by **TEMPORAL-PROJECTION.**

Causal Undercutting

CAUSAL-UNDERCUTTER

Where $t_0 \leq t_1$ and $(t_1+\varepsilon) < t$, "*A*-at-t_1 & *Q*-at-t_1 & (*A* when *Q* is causally sufficient for ~*P* after an interval ε)" is a defeasible undercutting defeater for the inference from *P*-at-t_0 to P-throughout-(t^*,t) by **TEMPORAL-PROJECTION.**

```
(def-backwards-undercutter CAUSAL-UNDERCUTTER
   :defeatee temporal-projection
   :forwards-premises
      "(A when Q is causally sufficient for ~P after an interval interval)"
      "(A at time1)"
        (:condition (and (time0 <= time1) ((time1 + interval) < time)))
   :backwards-premises
      "(Q at time1)"
   :variables  A Q P time0 time time* time1 interval op
   :defeasible?  T)
```

Causal Implication

- We want to use the causal connection to support inferences about what will happen.
  - "The gun is fired when the gun is loaded is causally sufficient for ~(Jones is alive) after an interval 20"does not imply that if the gun is fired at t and the gun is loaded at t then Jones is dead at t+20.
- All that is implied is that Jones is dead over some interval open on the left with t+20 as the lower bound.
  - We can conclude that *there is* at time > t+20 at which Jones is dead, but it does not follow as a matter of logic that Jones is dead at any particular time because, at least as far as this causal law is concerned, Jones could become alive again after becoming dead.
- To infer that Jones is dead at a particular time after t+20, we must combine the causal sufficiency with temporal projection.

Causal-Implication

CAUSAL-IMPLICATION

- If Q is temporally-projectible, $(t^{**}-(t+\varepsilon)) < \log(.5)/\log(\rho)$, and $((t+\varepsilon) \leq t^* < t^{**})$, then "($A$ when P is causally sufficient for Q after an interval ε) & A-at-t & P-at-t" is a defeasible reason for "Q-throughout-$(t^*, t^{**}]$" and for "Q-throughout-(t^*, t^{**})".
- If Q is temporally-projectible, $(t^{**}-(t+\varepsilon)) < \log(.5)/\log(\rho)$, and $((t+\varepsilon) < t^* \leq t^{**})$, then "($A$ when P is causally sufficient for Q after an interval ε) & A-at-t & P-at-t" is a defeasible reason for "Q-throughout-$[t^*, t^{**}]$".

```
(def-backwards-reason CAUSAL-IMPLICATION
   :conclusion  "(Q throughout (op time* time**))"
   :condition (and (<= time* time**)
                   ((time** - time*) < log(.5)/log(*temporal-decay*)))
   :forwards-premises
   "(A when P is causally sufficient for Q after an interval interval)"
     (:condition (every #'temporally-projectible (conjuncts Q)))
   "(A at time)"
   (:condition
    (or (and (eq op 'clopen) ((time + interval) <= time*) (time* < time**)
             ((time** - (time + interval)) < log(.5)/log(*temporal-decay*)))
        (and (eq op 'closed) ((time + interval) < time*) (time* <= time**)
             ((time** - (time + interval)) < log(.5)/log(*temporal-decay*)))
        (and (eq op 'open) ((time + interval) <= time*) (time* < time**)
             ((time** - (time + interval)) < log(.5)/log(*temporal-decay*)))))
   :backwards-premises
   "(P at time)"
   :variables  A P Q interval time time* time** op
   :strength  (- (* 2 (expt *temporal-decay* (- time** time))) 1)
   :defeasible?  T)
```

The Yale Shooting Problem

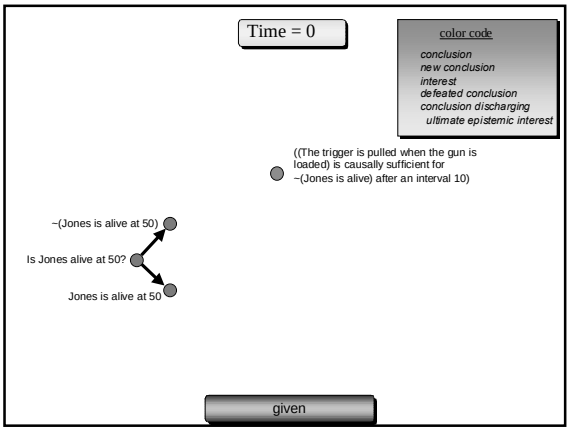
I know that the gun being fired while loaded will cause Jones to become dead.

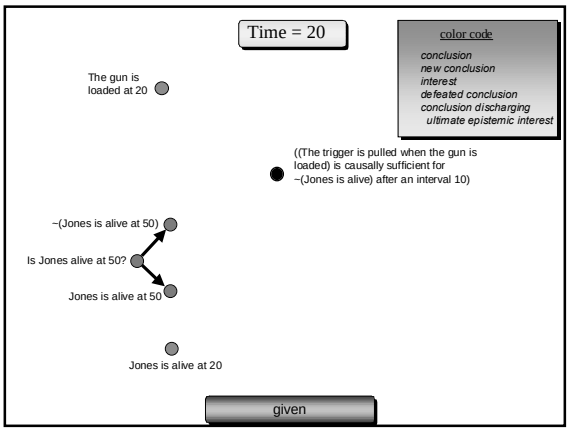
I know that the gun is initially loaded, and Jones is initially alive.

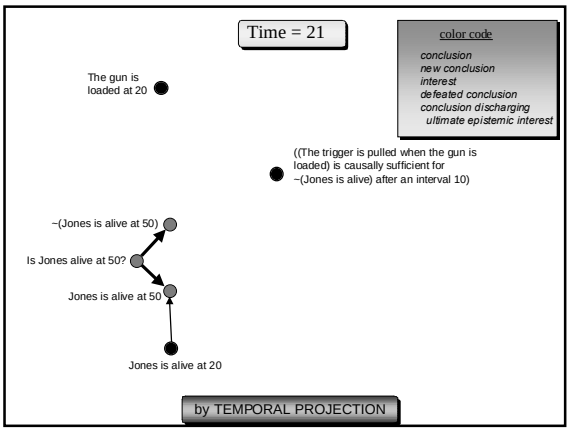
Later, the gun is fired.

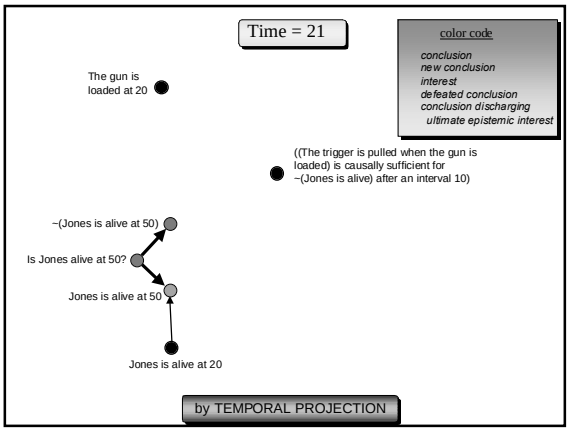
Should I conclude that Jones becomes dead?

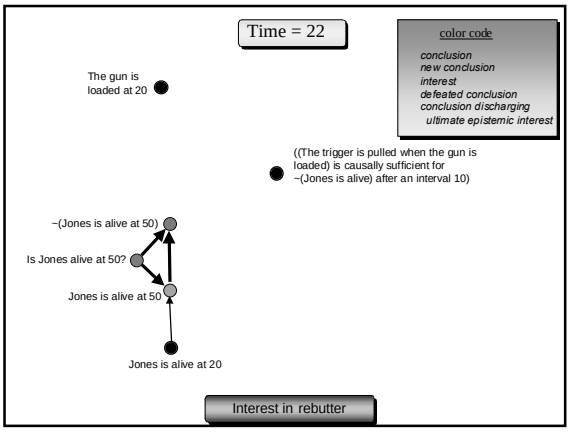
see OSCAR do it

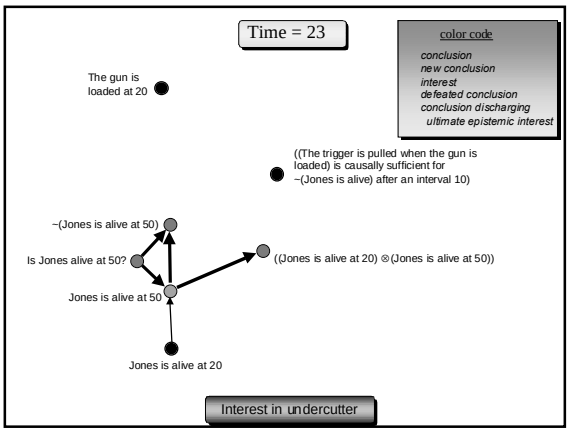


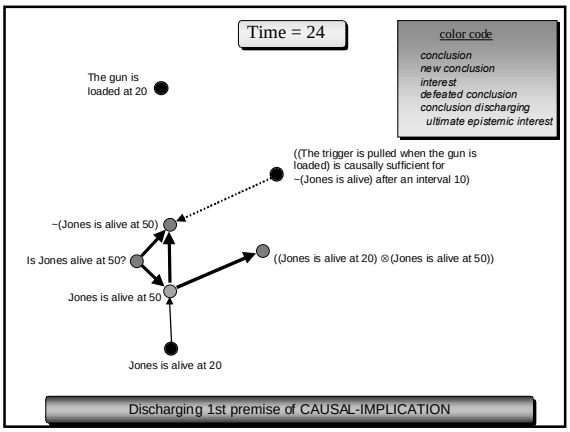


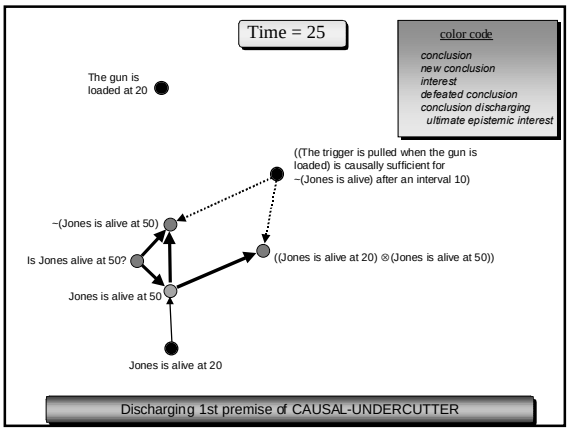


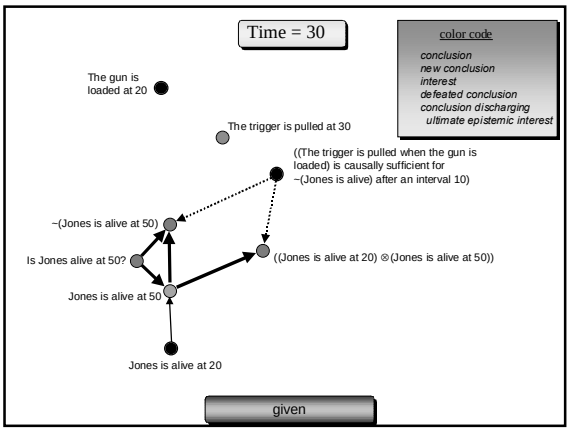

Time = 21
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
Jones is alive at 50
Jones is alive at 20
by TEMPORAL PROJECTION

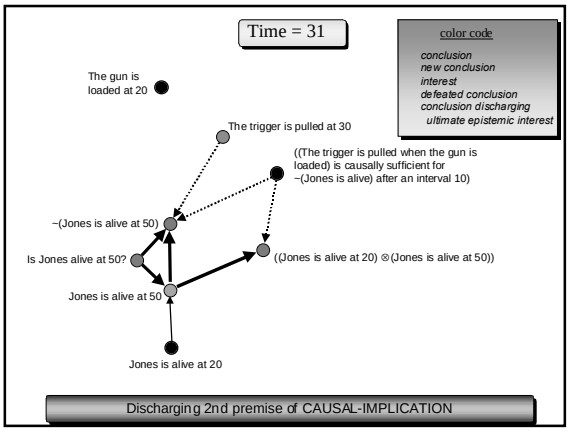

Time = 22
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
Jones is alive at 50
Jones is alive at 20
Interest in rebutter

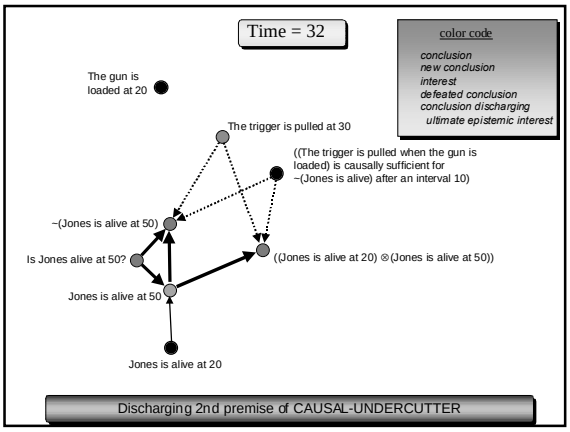

Time = 23
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in undercutter

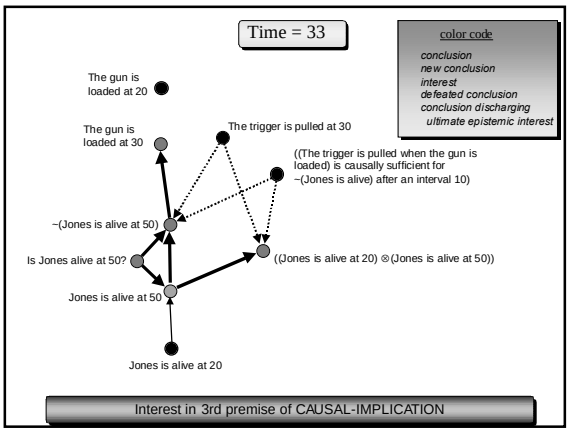

Time = 24
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 1st premise of CAUSAL-IMPLICATION

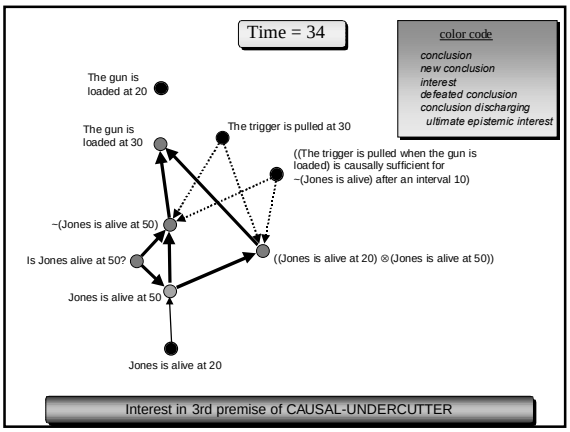

Time = 25
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 1st premise of CAUSAL-UNDERCUTTER

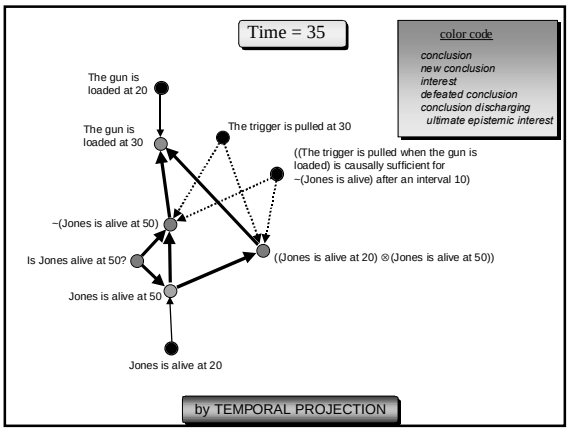

Time = 30
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
given

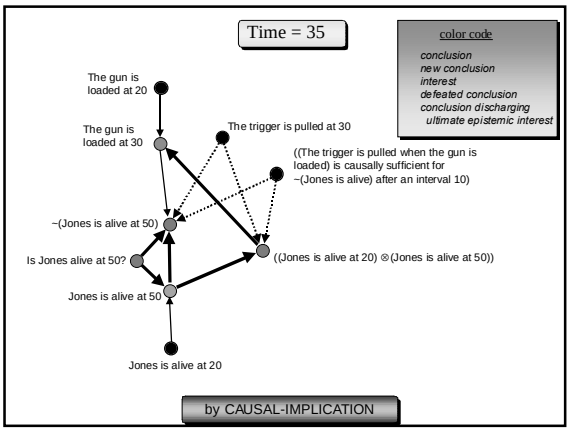

Time = 31
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 2nd premise of CAUSAL-IMPLICATION

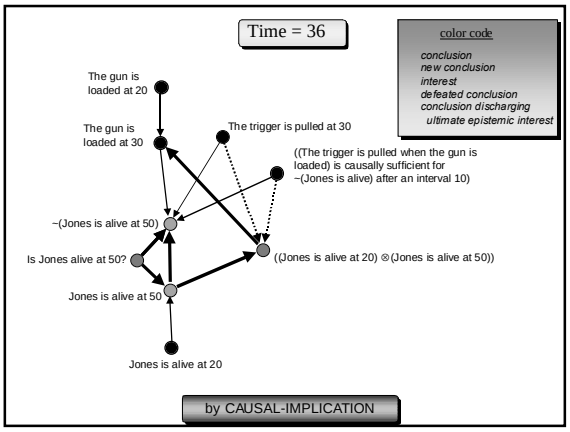

Time = 32
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 2nd premise of CAUSAL-UNDERCUTTER

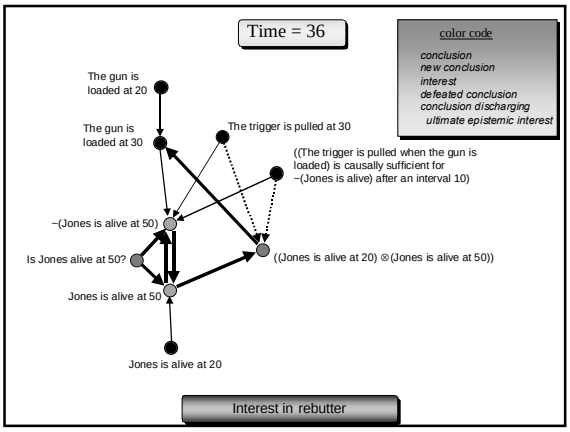

Time = 33
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in 3rd premise of CAUSAL-IMPLICATION

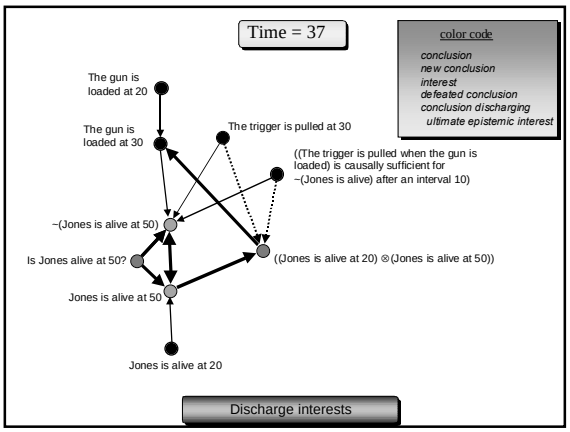

Time = 34
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in 3rd premise of CAUSAL-UNDERCUTTER

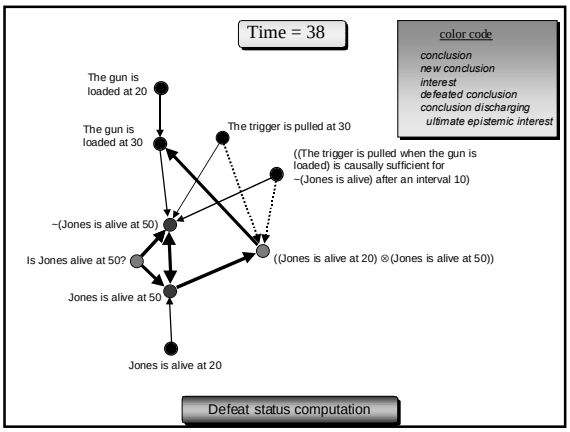

Time = 35
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
by TEMPORAL PROJECTION

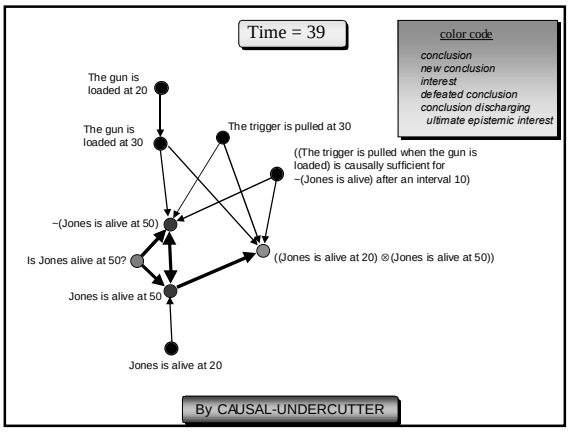

Time = 35
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
by CAUSAL-IMPLICATION

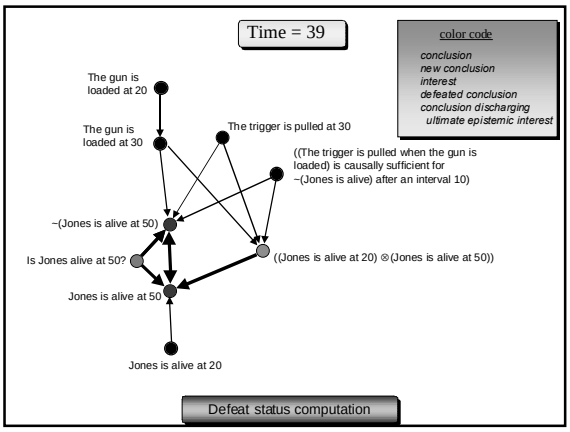

Time = 36
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
by CAUSAL-IMPLICATION

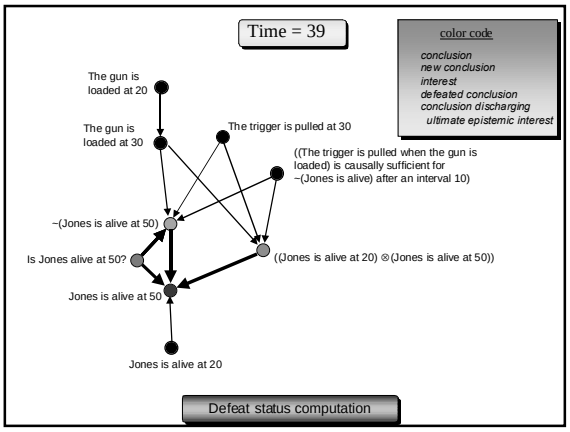

Time = 36
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in rebutter

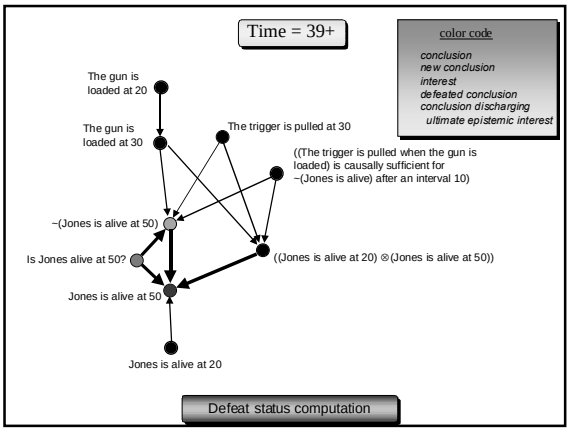

Time = 37
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharge interests

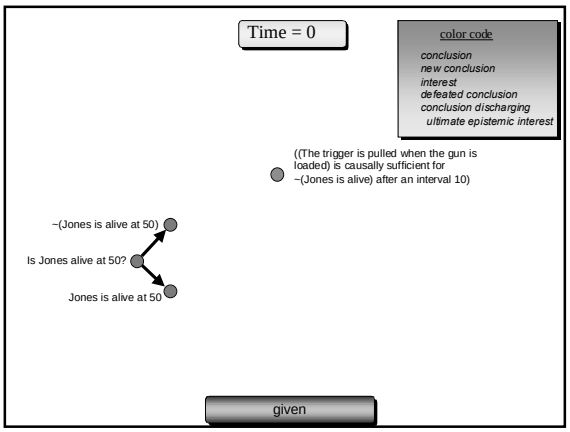

Time = 38
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation

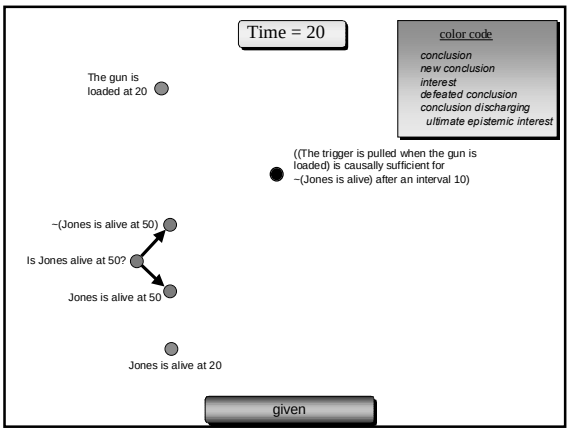

Time = 39
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
By CAUSAL-UNDERCUTTER

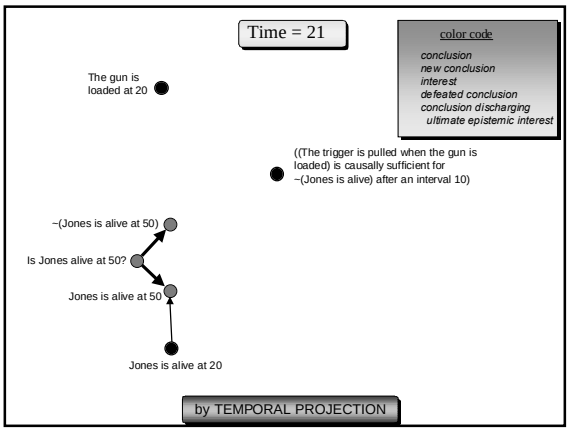

Time = 39
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation

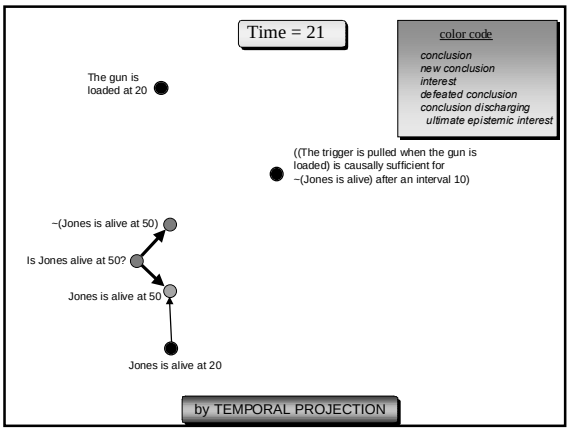

Time = 39
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation

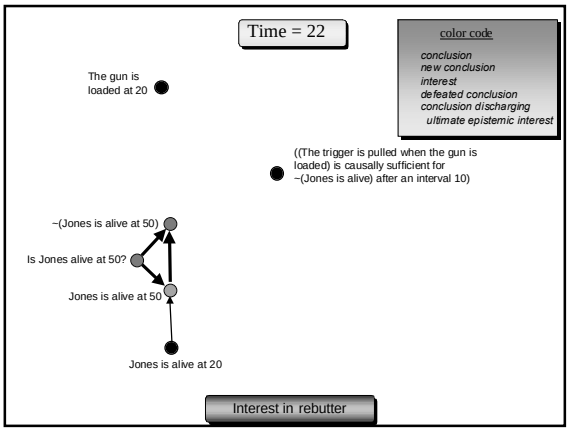

Time = 39+
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation

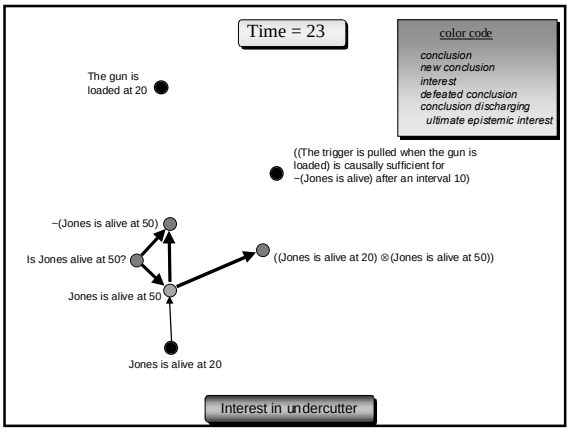

Time = 0
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
Jones is alive at 50
given

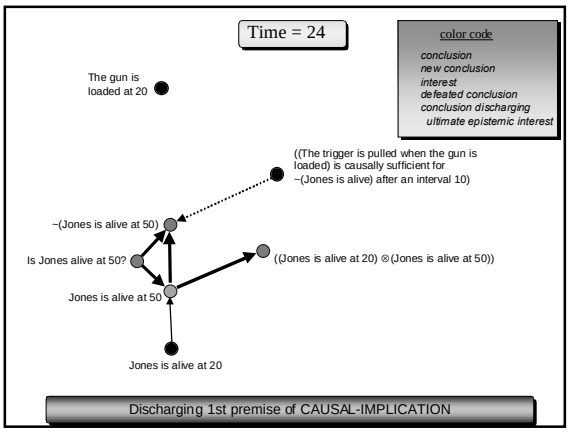

Time = 20
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
Jones is alive at 50
Jones is alive at 20
given

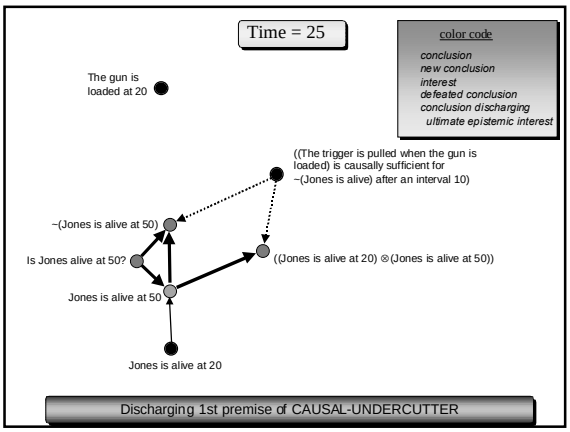

Time = 21
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
Jones is alive at 50
Jones is alive at 20
by TEMPORAL PROJECTION

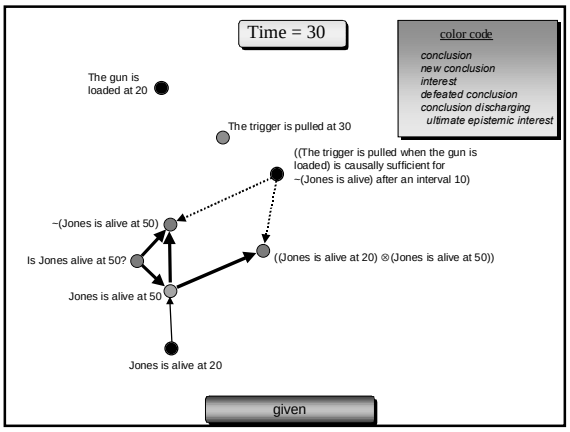

Time = 21
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
Jones is alive at 50
Jones is alive at 20
by TEMPORAL PROJECTION

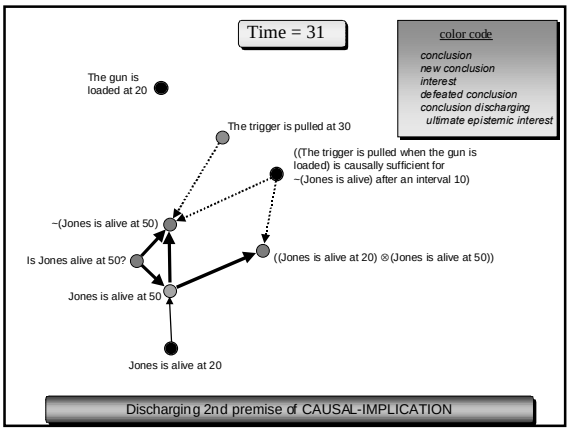

Time = 22
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
Jones is alive at 50
Jones is alive at 20
Interest in rebutter

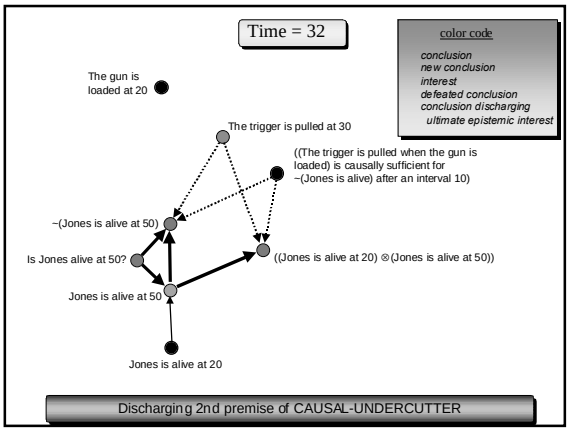

Time = 23
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in undercutter

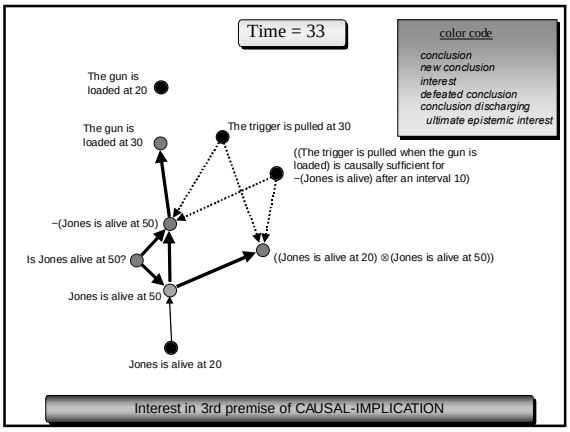

Time = 24
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 1st premise of CAUSAL-IMPLICATION

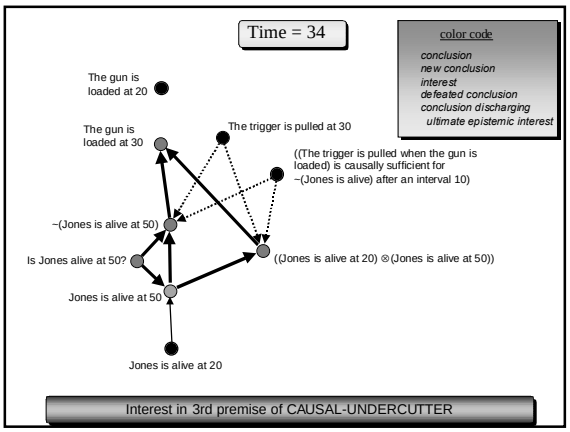

Time = 25
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 1st premise of CAUSAL-UNDERCUTTER

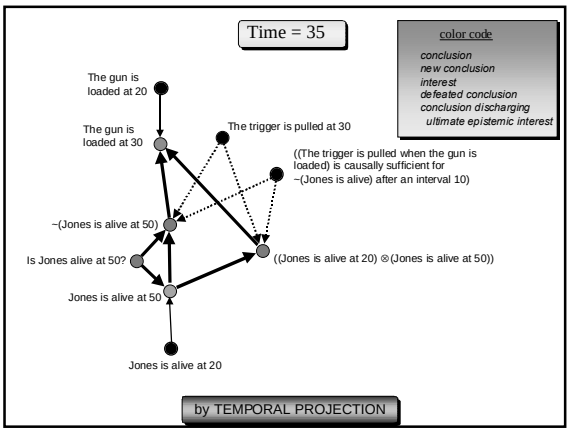

Time = 30
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
given

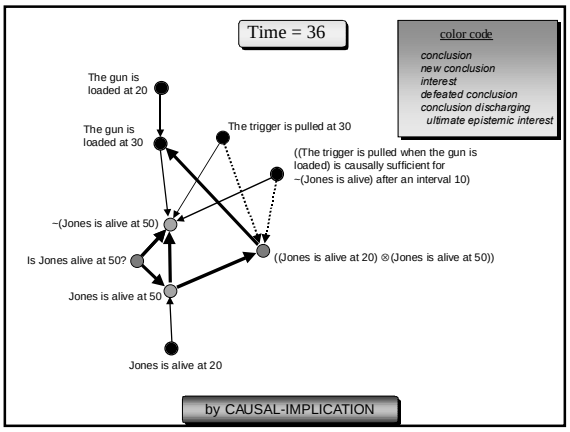

Time = 31
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 2nd premise of CAUSAL-IMPLICATION

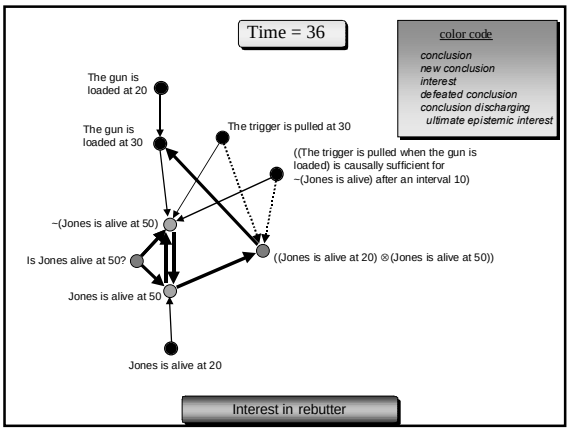

Time = 32
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharging 2nd premise of CAUSAL-UNDERCUTTER

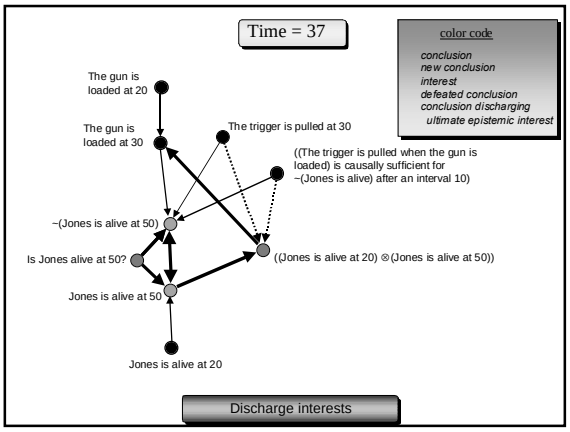

Time = 33
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in 3rd premise of CAUSAL-IMPLICATION

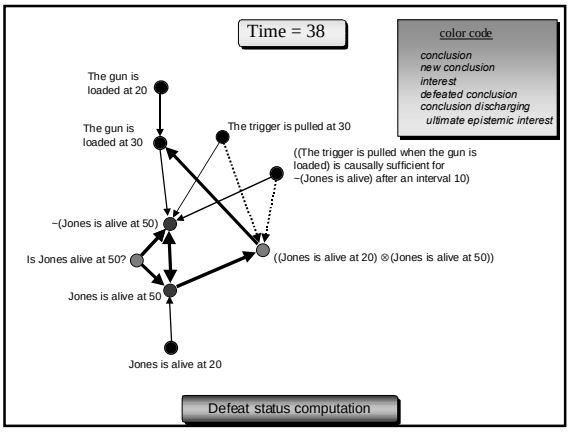

Time = 34
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in 3rd premise of CAUSAL-UNDERCUTTER

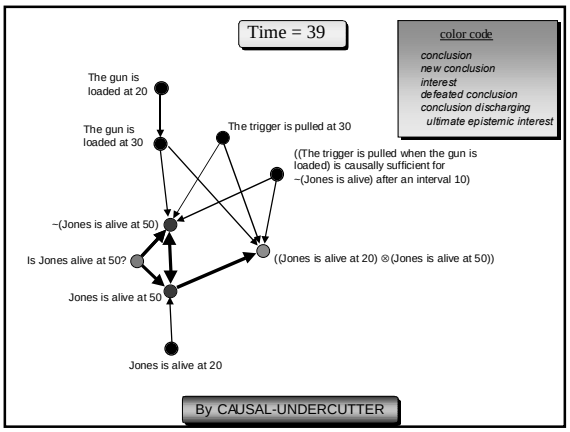

Time = 35
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
by TEMPORAL PROJECTION

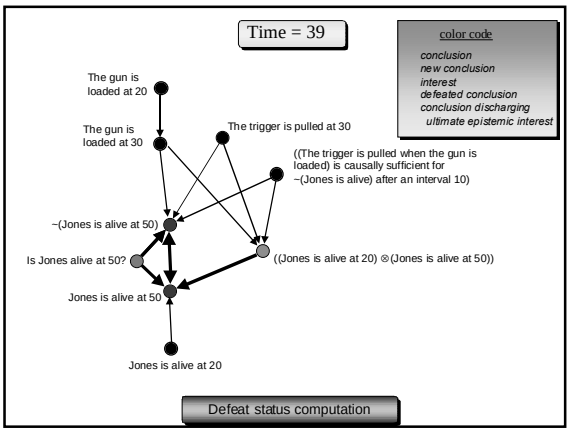

Time = 36
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
by CAUSAL-IMPLICATION

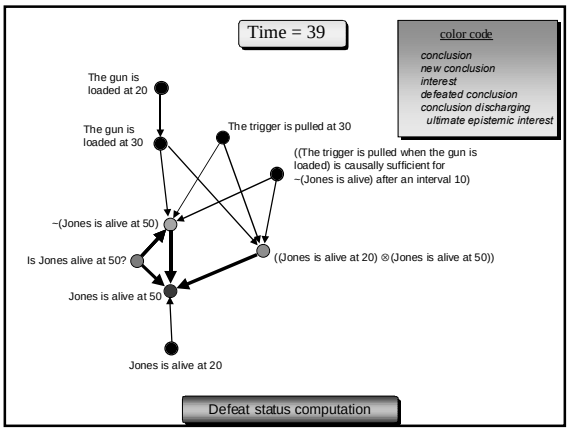

Time = 36
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Interest in rebutter

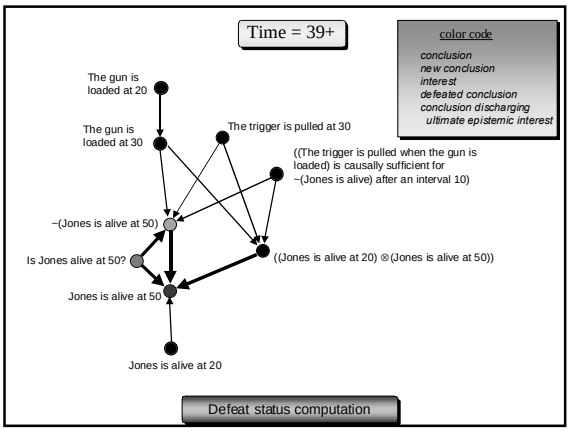

Time = 37
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Discharge interests


Time = 38
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation


Time = 39
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
By CAUSAL-UNDERCUTTER


Time = 39
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation


Time = 39
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation


Time = 39+
color code
conclusion
new conclusion
interest
defeated conclusion
conclusion discharging
ultimate epistemic interest
The gun is loaded at 20
The gun is loaded at 30
The trigger is pulled at 30
((The trigger is pulled when the gun is loaded) is causally sufficient for ~(Jones is alive) after an interval 10)
~(Jones is alive at 50)
Is Jones alive at 50?
((Jones is alive at 20) ⊗(Jones is alive at 50))
Jones is alive at 50
Jones is alive at 20
Defeat status computation

The Qualification Problem

- The Frame Problem concerned the proliferation of frame axioms—axioms concerning what does not change.
- The Qualification Problem concerned the difficulty in correctly formulating axioms about what does change.
- The Qualification Problem is the problem of getting the antecedent right in axioms like "A match's being struck when it is dry, in the presence of oxygen, ... , is causally sufficient for it to light".
- The difficulty is that we are typically unable to fill in the ellipsis and give a complete list of the conditions required to cause a particular effect.

The Qualification Problem

- Within the present framework, the solution to the Qualification Problem seems to be fairly simple.\
  A when P is causally sufficient for Q after an interval ε" means
  $(\forall t)\{(A\text{-at-}t \ \&\ P\text{-at-}t) \Rightarrow (\exists\delta)Q\text{-throughout-}(t{+}\varepsilon\ ,\ t{+}\varepsilon{+}\delta]\}$.
- The causal knowledge that we use in reasoning about change is not generally of this form.
  - First, we rarely have more than a rough estimate of the value of ε.
  - Second, we are rarely in a position to formulate P precisely.

The Qualification Problem

- Our knowledge actually takes the form:
  $(\exists P^*)(\exists\varepsilon*)[P^*$ is true & $\varepsilon* \leq \varepsilon$ & (A when (P & P^*) is causally sufficient for Q after an interval $\varepsilon*$)].
- P formulates the known preconditions for the causal sufficiency, P* the unknown precondtions, and ε is the known upper bound on ε*.
- Let us abbreviate this as "A when P is weakly causally sufficient for Q after an interval ε".
- We acquire knowledge of weak causal sufficiency inductively.
  - For example, we learn inductively that striking a dry match is usually weakly causally sufficient for it to light after a negligible interval.

The Qualification Problem

- CAUSAL-UNDERCUTTER and CAUSAL-IMPLICATION both continue to hold if we reconstrue "causally suffficient" to mean "weakly causally sufficient".
- Thus we can reason about change in the same way even with incomplete causal knowledge.
- This resolves the Qualification Problem.

The Ramification Problem

- The Ramification Problem arises from the observation that in realistically complex environments, we cannot formulate axioms that completely specify the effects of actions or events.
- People sometimes refer to these as "actions with ramifications", as if these were peculiar actions.
- But in the real world, all actions have infinitely many ramifications stretching into the indefinite future.
- This is a problem for reasoning about change deductively, but does not seem to be a problem for reasoning about change defeasibly in the present framework.

The Ramification Problem

- Reasoning in OSCAR is interest-driven, and CAUSAL-IMPLICATION is a backwards-reason.
- This means that we only reason about potential effects of actions and events when they are of interest to us.
- Whether they are of interest can vary from circumstance to circumstance, allowing our reasoning to vary similarly, without our having to revise our knowledge base or rules of inference to accommodate the change.
- Deductive reasoning in terms of successor-state axioms is too crude an implementation to reflect this feature of human reasoning, but the current framework handles it automatically.
- The conclusion is that the Ramification Problem simply does not arise in this framework.