“Belief Revision” and Truth-Finding

Kevin T. Kelly
Department of Philosophy
Carnegie Mellon University
kk3n@andrew.cmu.edu

Further Reading

(with O. Schulte and V. Hendricks) "Reliable Belief Revision", in *Logic and Scientific Methods*, Dordrecht: Kluwer, 1997.

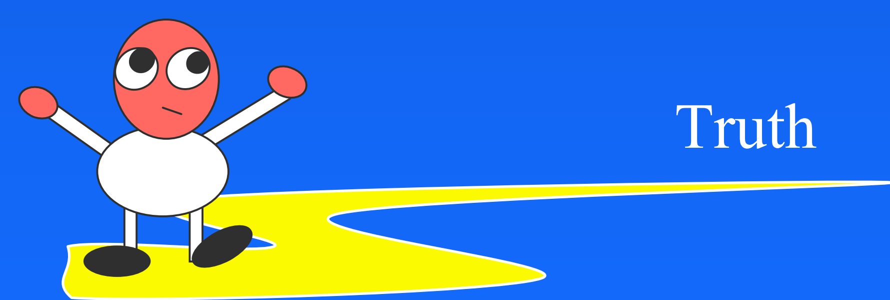
"The Learning Power of Iterated Belief Revision", in *Proceedings of the Seventh TARK Conference*, 1998.

"Iterated Belief Revision, Reliability, and Inductive Amnesia," *Erkenntnis*, 50: 1998

The Idea

- Belief revision theory... "*rational*" belief change

- Learning theory...............*reliable* belief change
- **Conflict?**



Part I

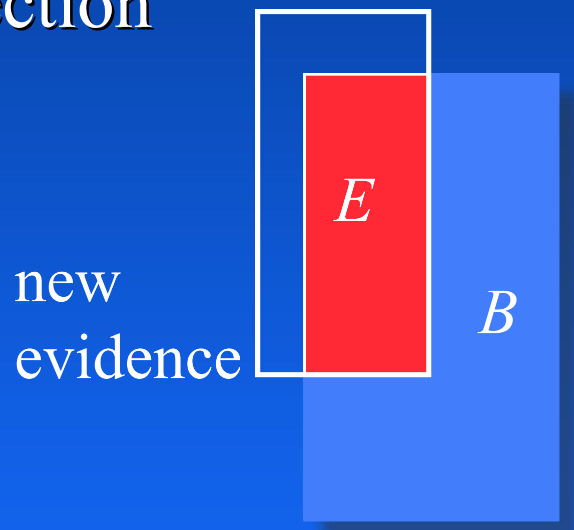
Iterated Belief Revision

Bayesian (Vanilla) Updating

Propositional epistemic state

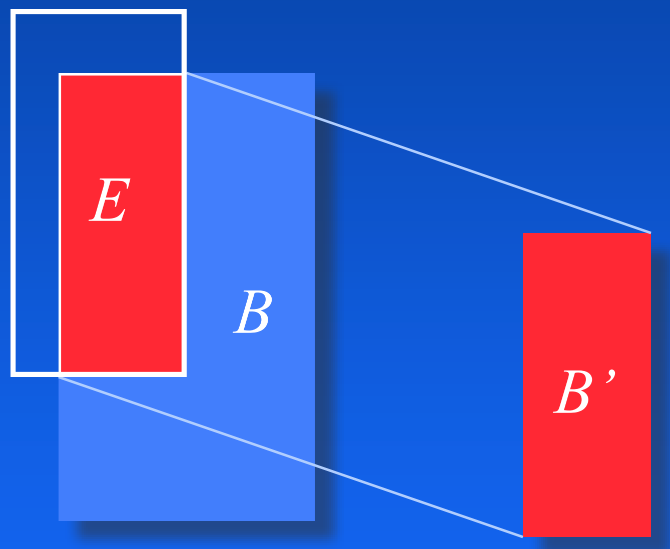
Bayesian (Vanilla) Updating

- New belief is intersection
- Perfect memory
- No inductive leaps



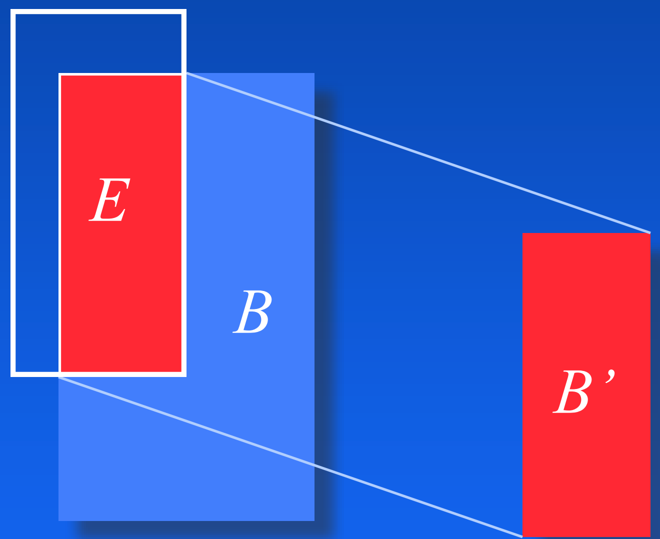
Bayesian (Vanilla) Updating

- New belief is intersection
- Perfect memory
- No inductive leaps

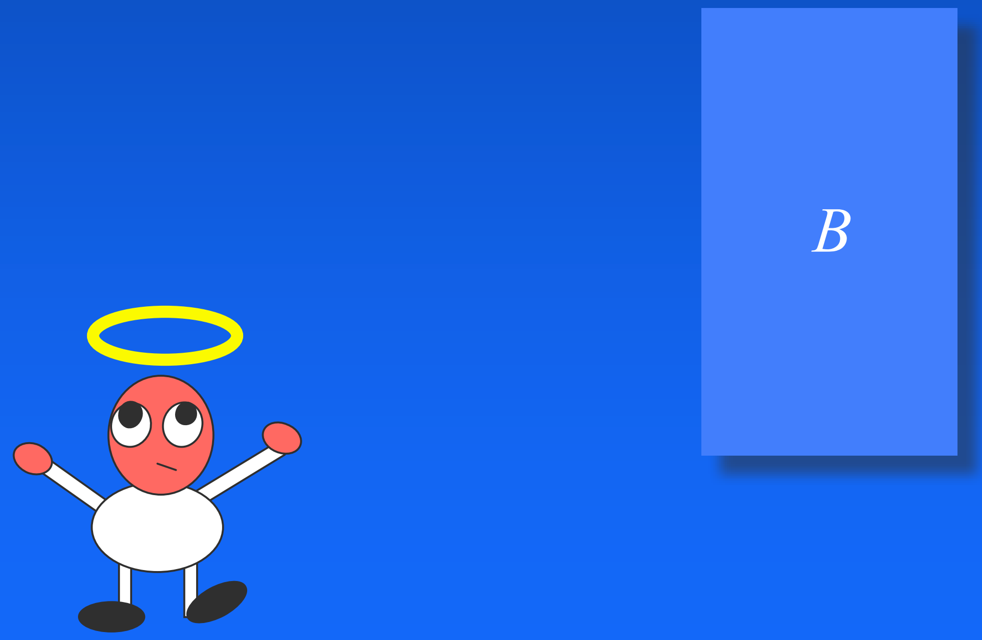


Bayesian (Vanilla) Updating

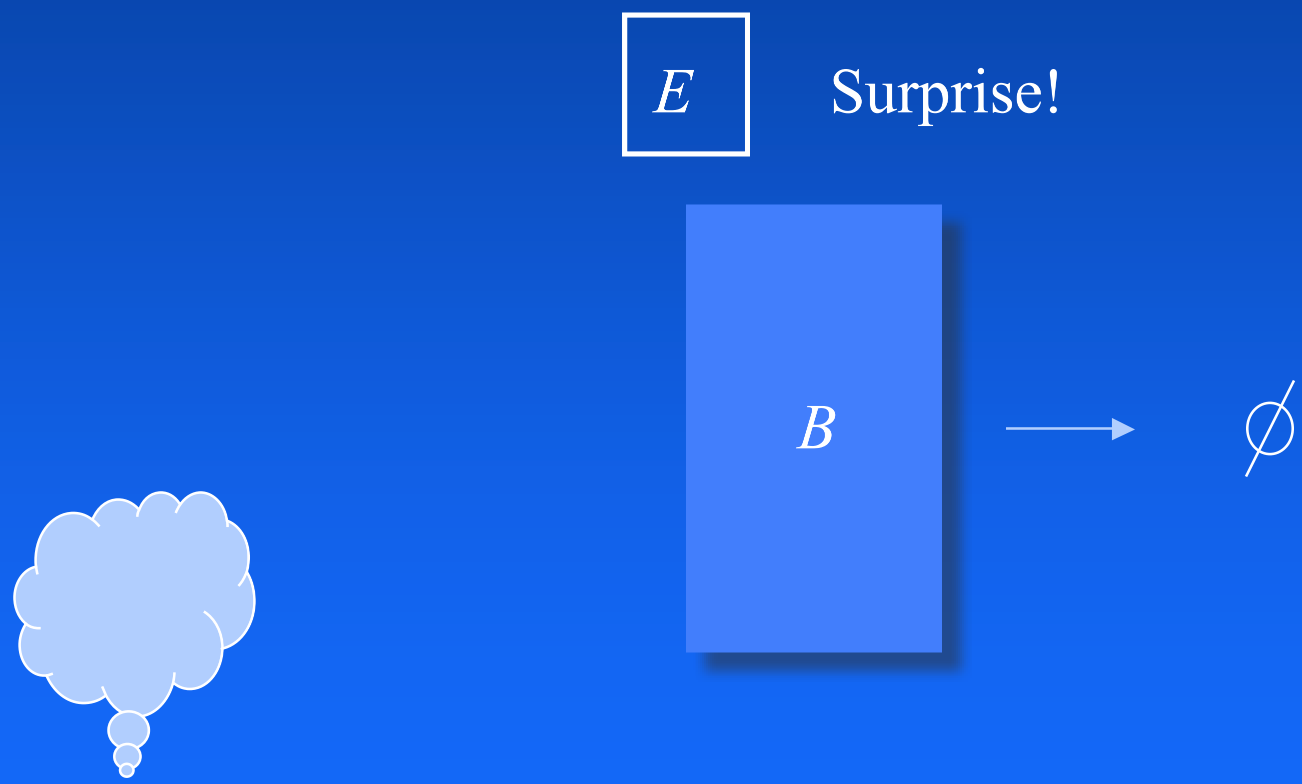
- New belief is intersection
- Perfect memory
- No inductive leaps



“Epistemic Hell” (a.k.a. Nirvana)

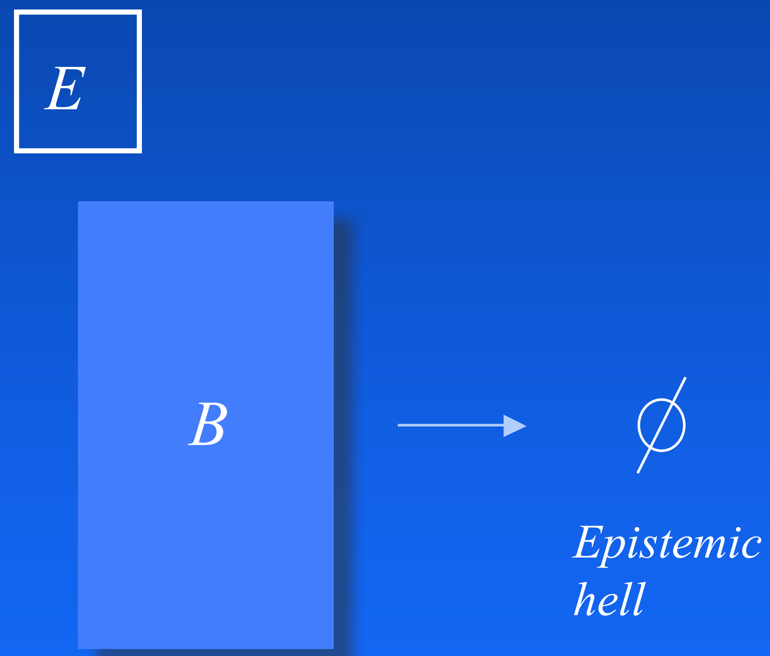


"Epistemic Hell" (a.k.a. Nirvana)



Epistemic Hell (a.k.a. Nirvana)

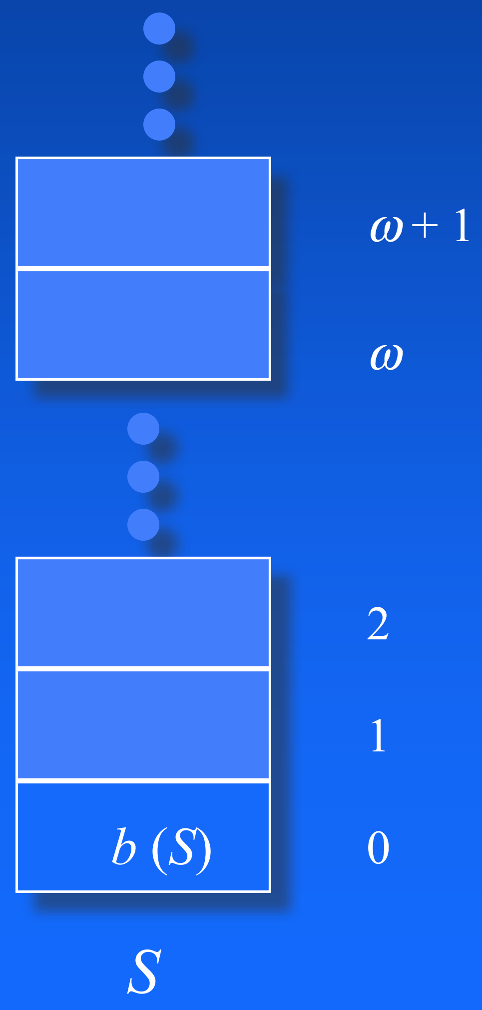
- Scientific revolutions
- Suppositional reasoning
- Conditional pragmatics
- Decision theory
- Game theory
- Data bases



Ordinal Epistemic States

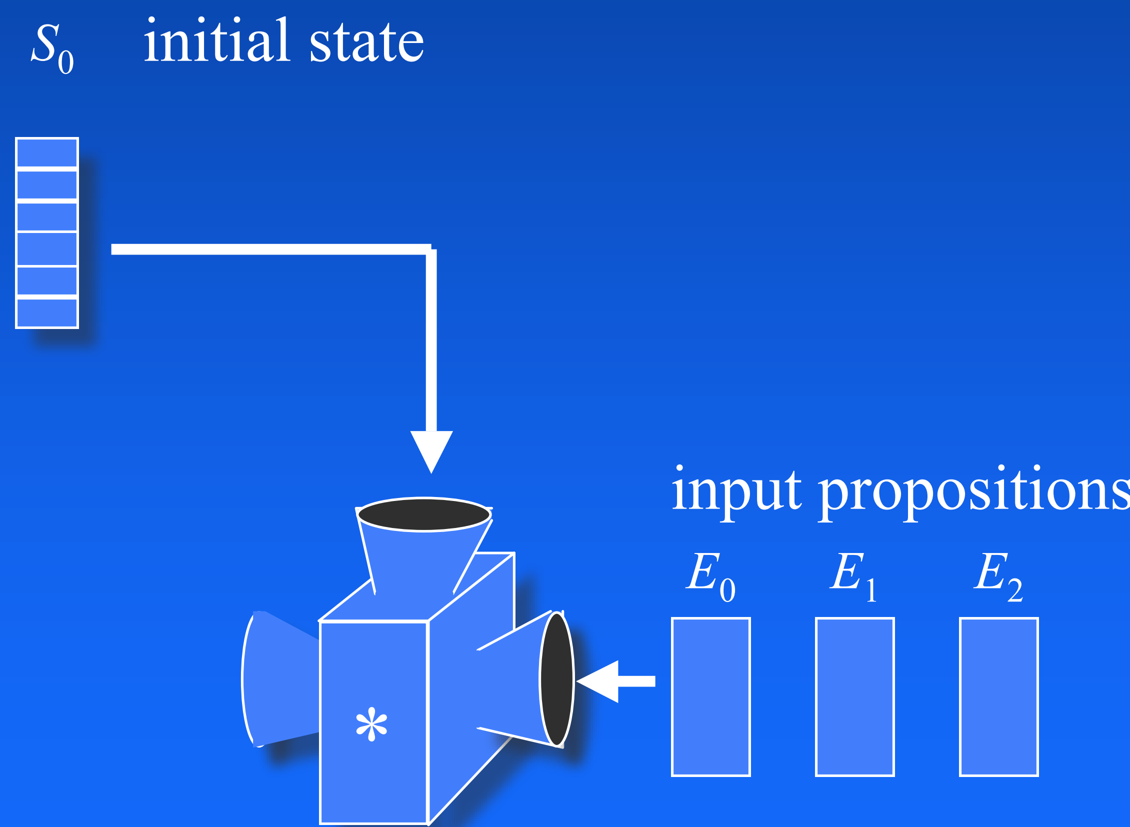
Spohn 88

- Ordinal-valued degrees of "implausibility"
- Belief state is bottom level

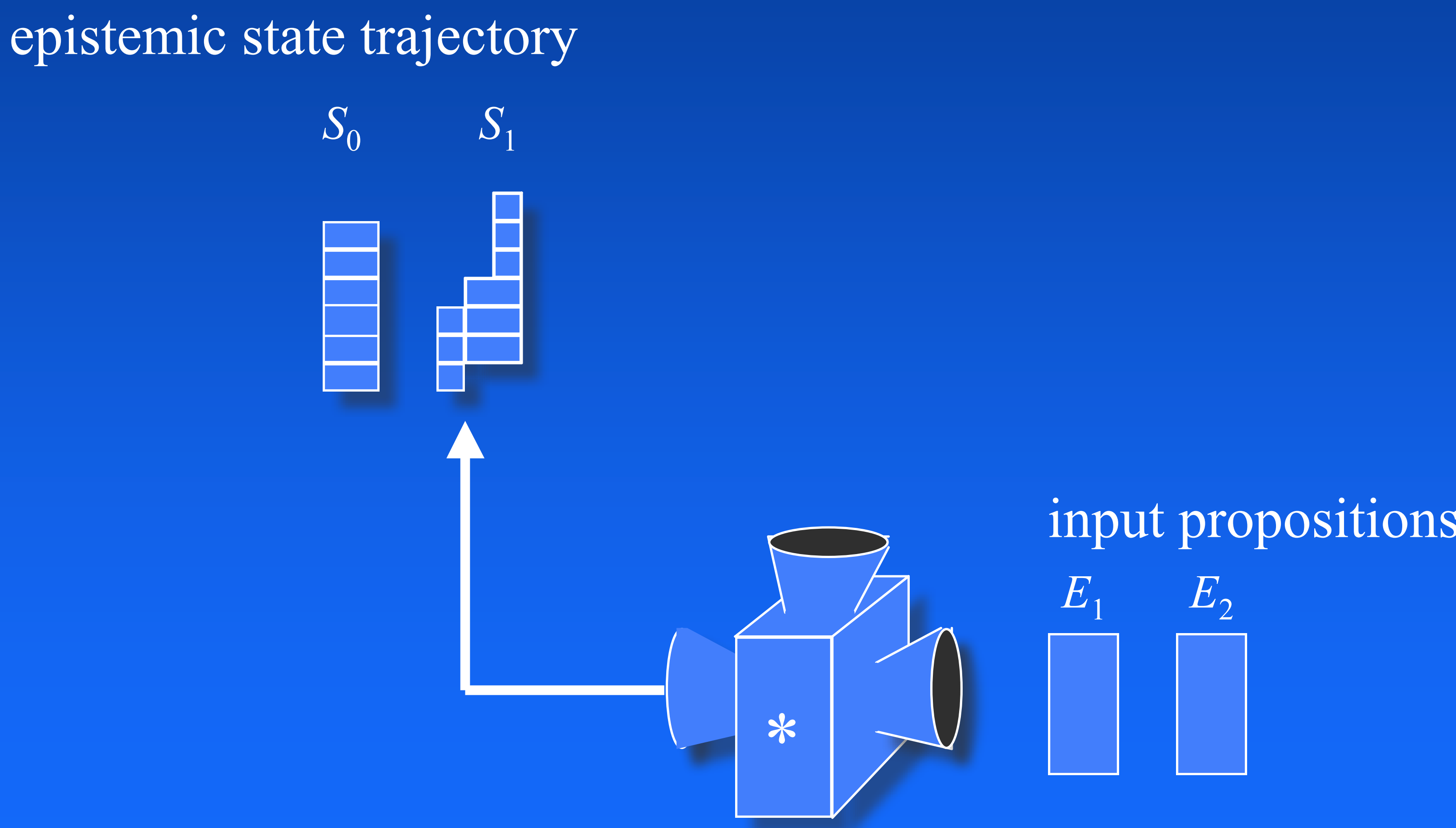


Iterated Belief Revision

epistemic state trajectory

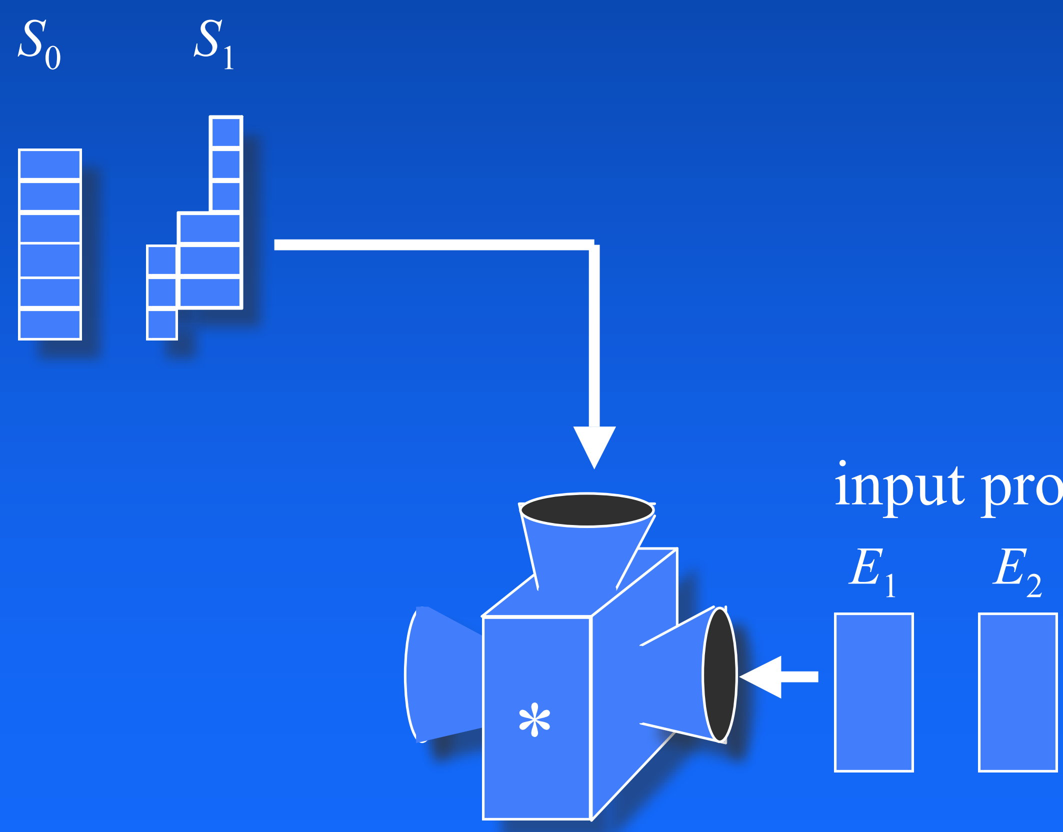


Iterated Belief Revision

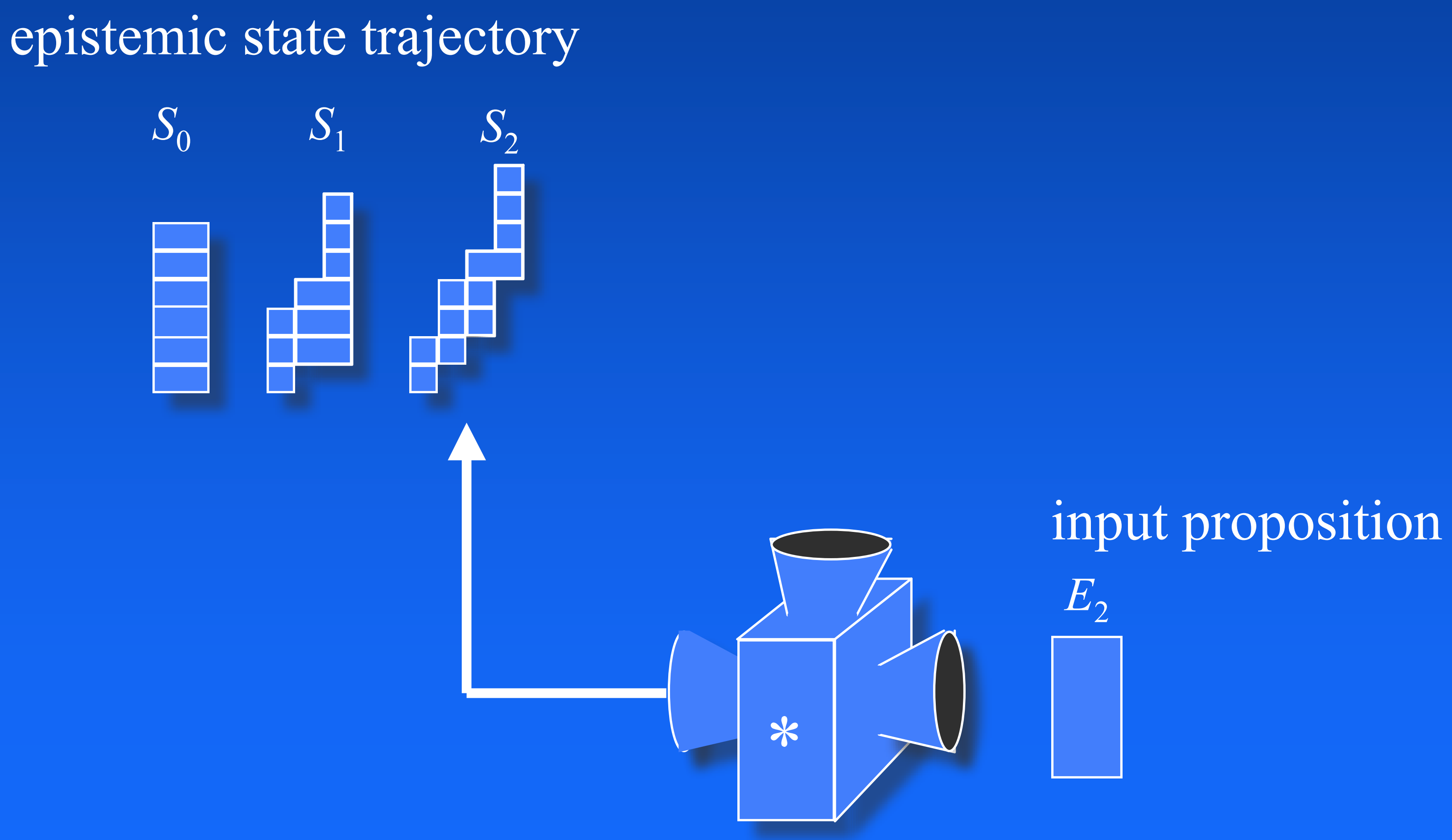


Iterated Belief Revision

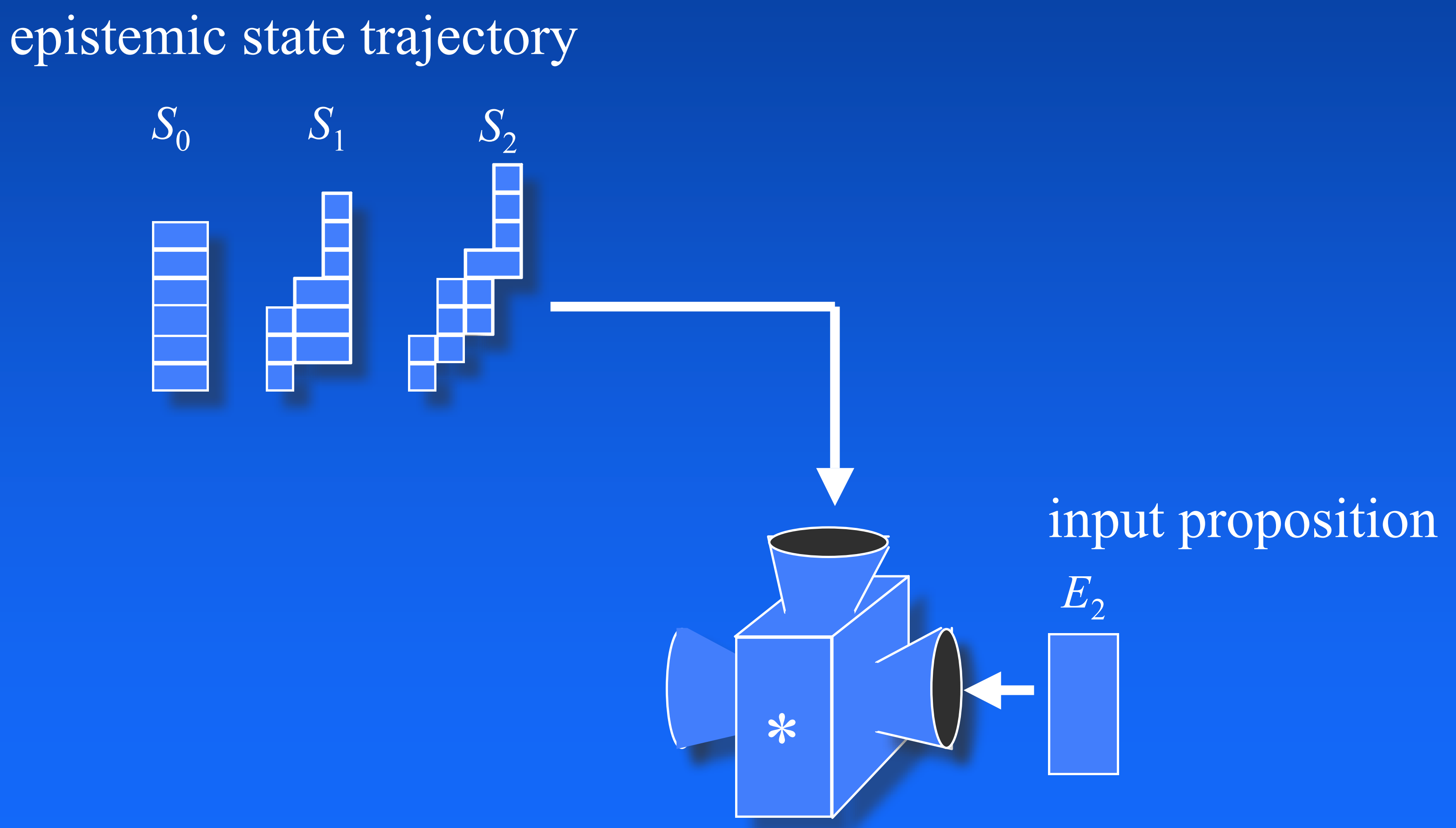
epistemic state trajectory



Iterated Belief Revision

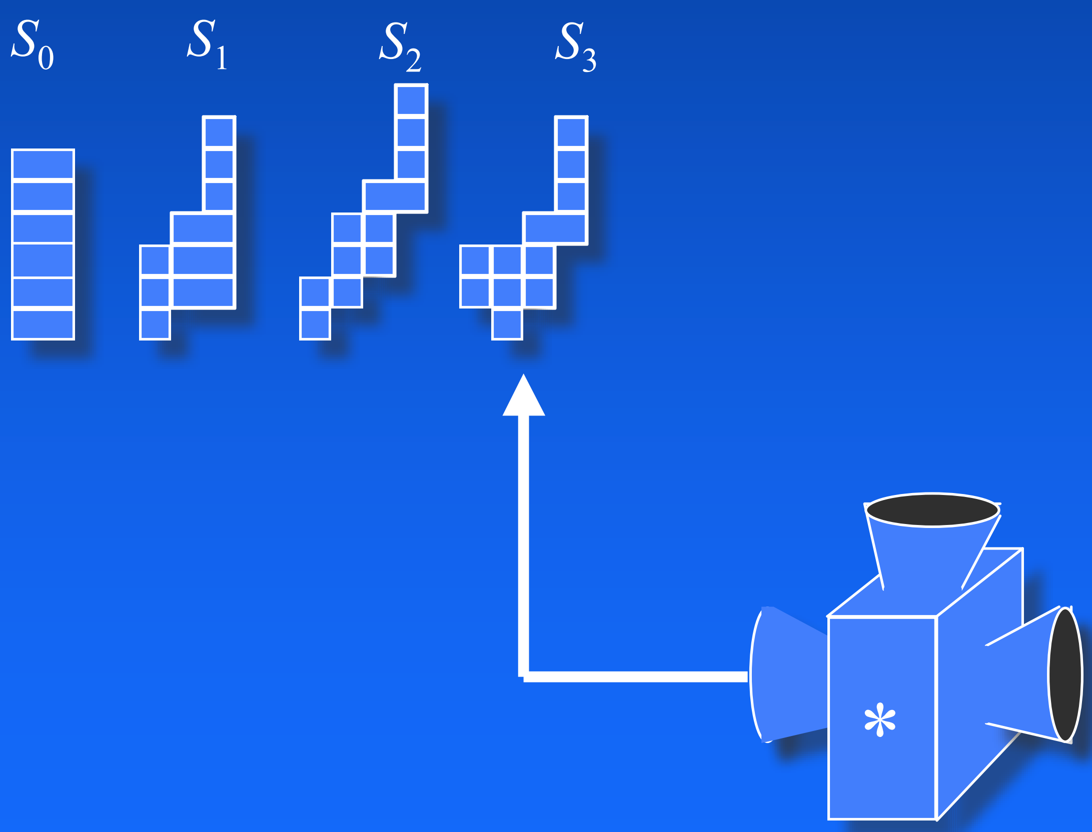


Iterated Belief Revision



Iterated Belief Revision

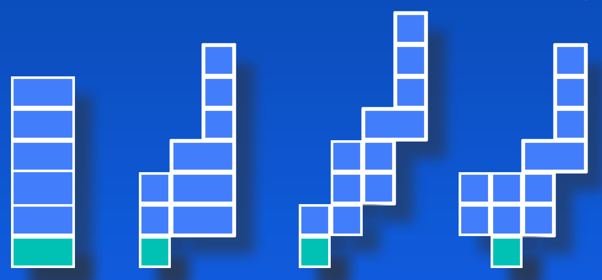
epistemic state trajectory



Iterated Belief Revision

epistemic state trajectory

S_0 S_1 S_2 S_3

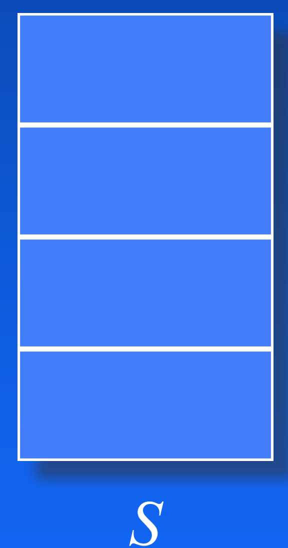


$b\ (S_0)$ $b\ (S_1)$ $b\ (S_2)$ $b\ (S_3)$

belief state trajectory

Generalized Conditioning *C

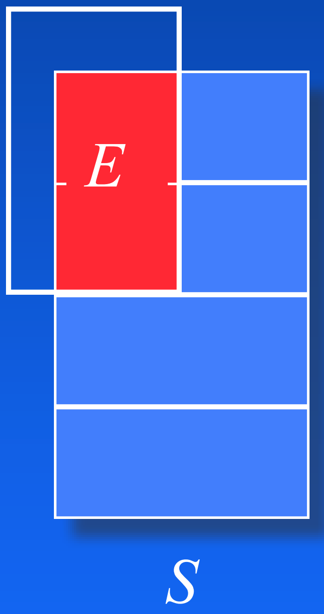
Spohn 88



Generalized Conditioning *C

Spohn 88

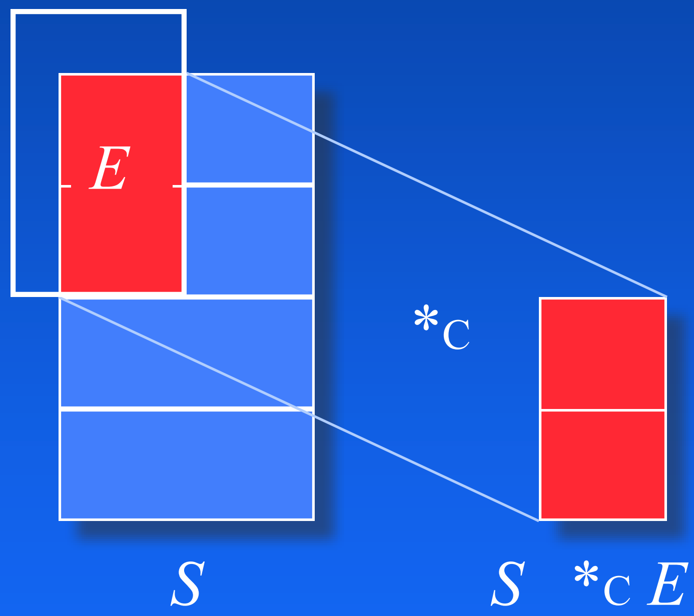
- Condition entire epistemic state



Generalized Conditioning $*_C$

Spohn 88

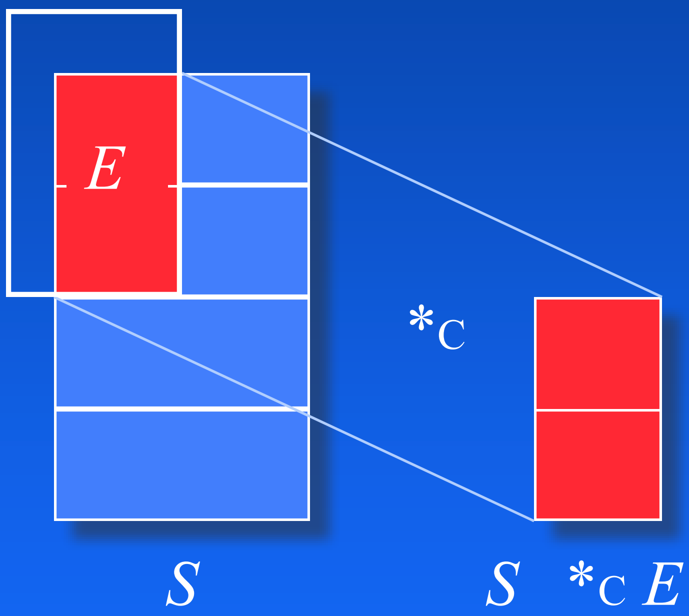
- Condition entire epistemic state



Generalized Conditioning $*_C$

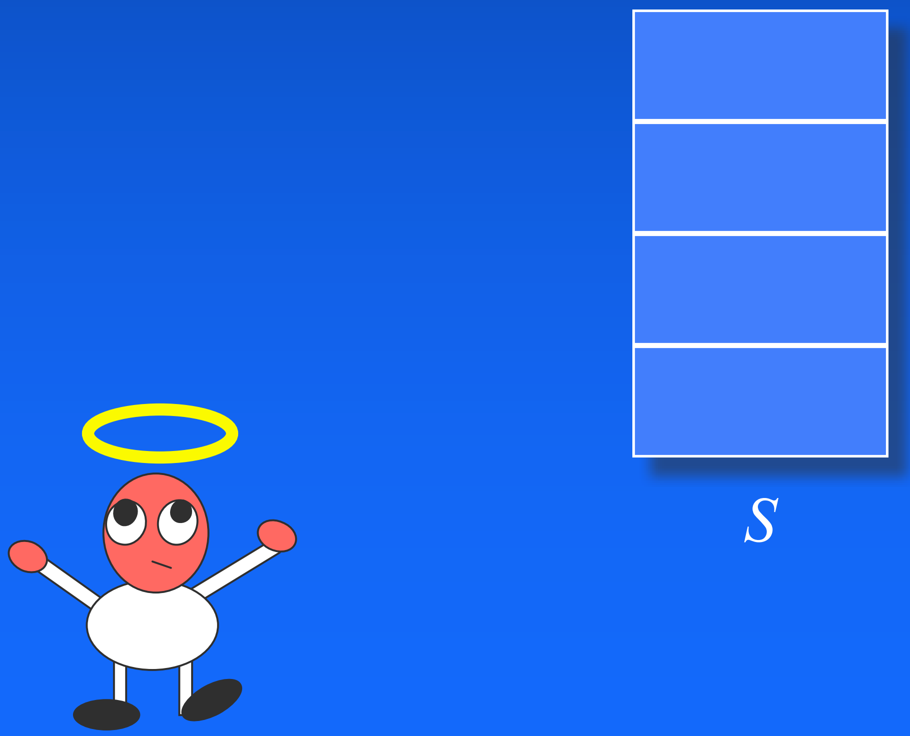
Spohn 88

- Condition entire epistemic state
- Perfect memory
- Inductive leaps
- No epistemic hell *if* evidence sequence is consistent



Lexicographic Updating *L

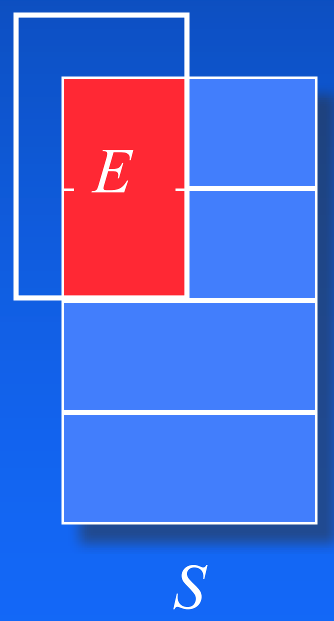
Spohn 88, Nayak 94



Lexicographic Updating $*_L$

Spohn 88, Nayak 94

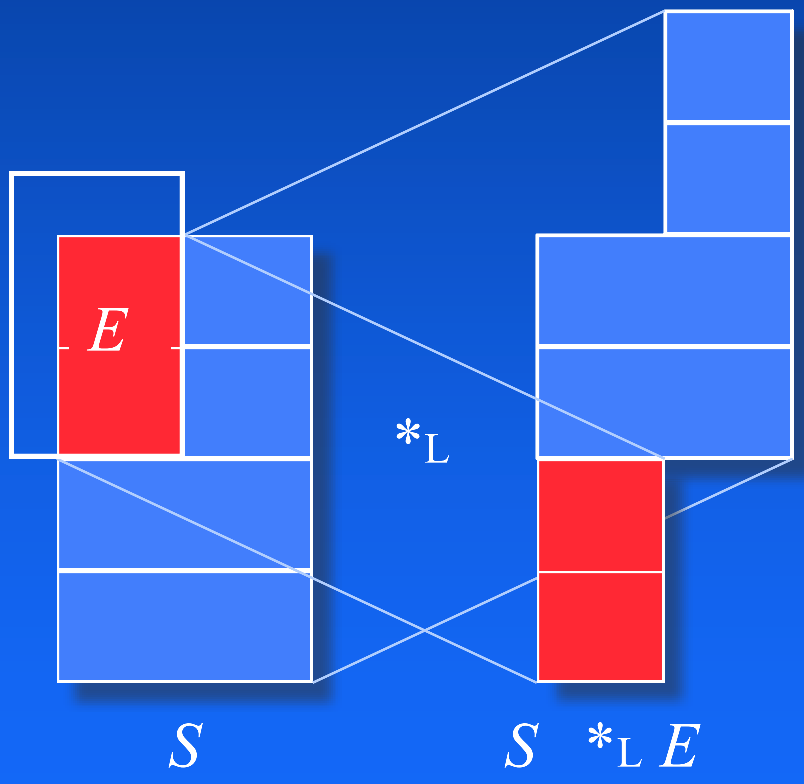
- Lift refuted possibilities above non-refuted possibilities preserving order.



Lexicographic Updating $*_L$

Spohn 88, Nayak 94

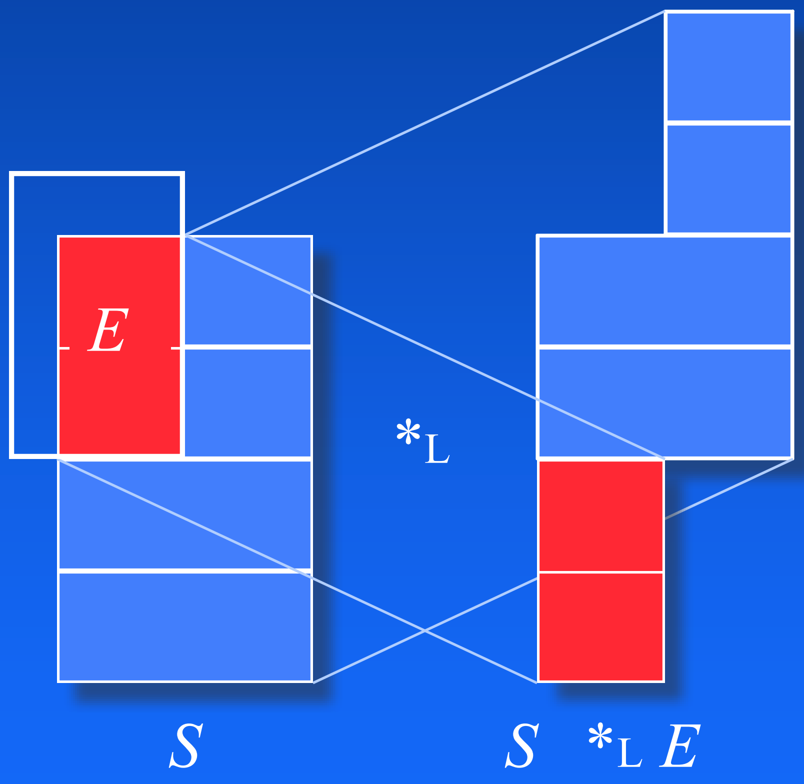
- Lift refuted possibilities above non-refuted possibilities preserving order.



Lexicographic Updating $*_L$

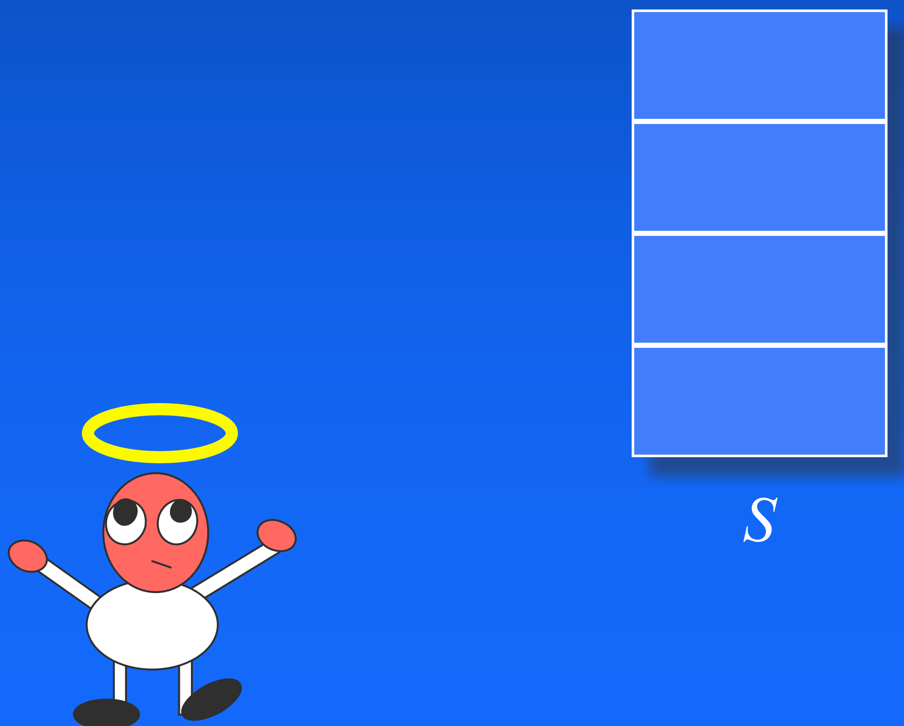
Spohn 88, Nayak 94

- Lift refuted possibilities above non-refuted possibilities preserving order.
- Perfect memory on consistent data sequences
- Inductive leaps
- No epistemic hell



Minimal or "Natural" Updating $*_M$

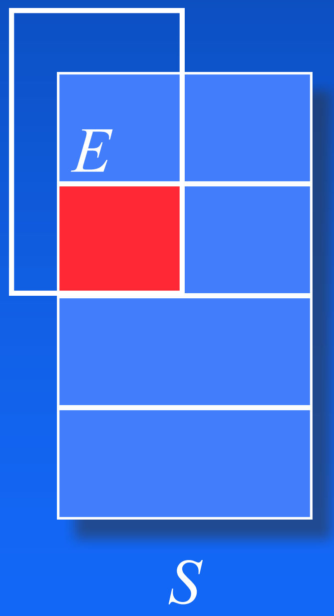
Spohn 88, Boutilier 93



Minimal or "Natural" Updating $*_M$

Spohn 88, Boutilier 93

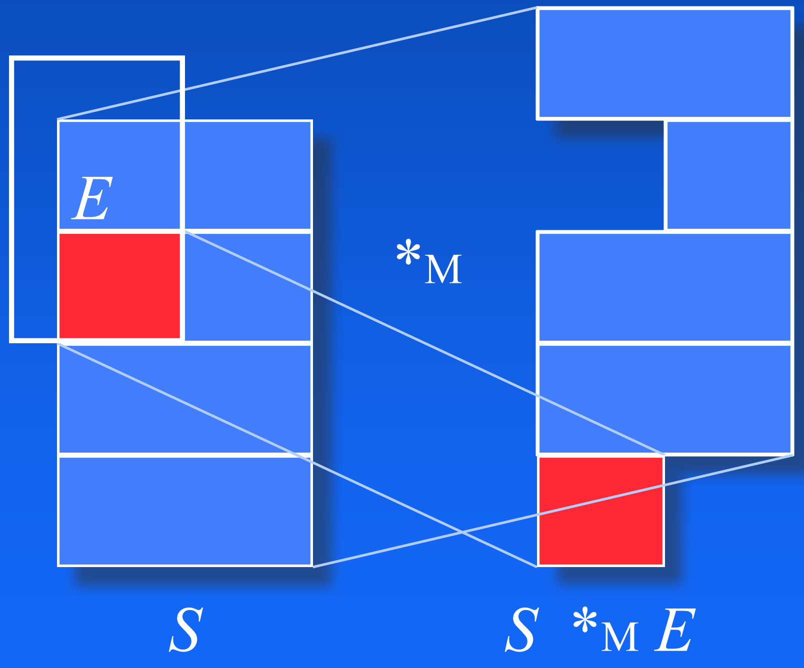
- Drop the lowest possibilities consistent with the data to the bottom and raise everything else up one notch



Minimal or "Natural" Updating $*_M$

Spohn 88, Boutilier 93

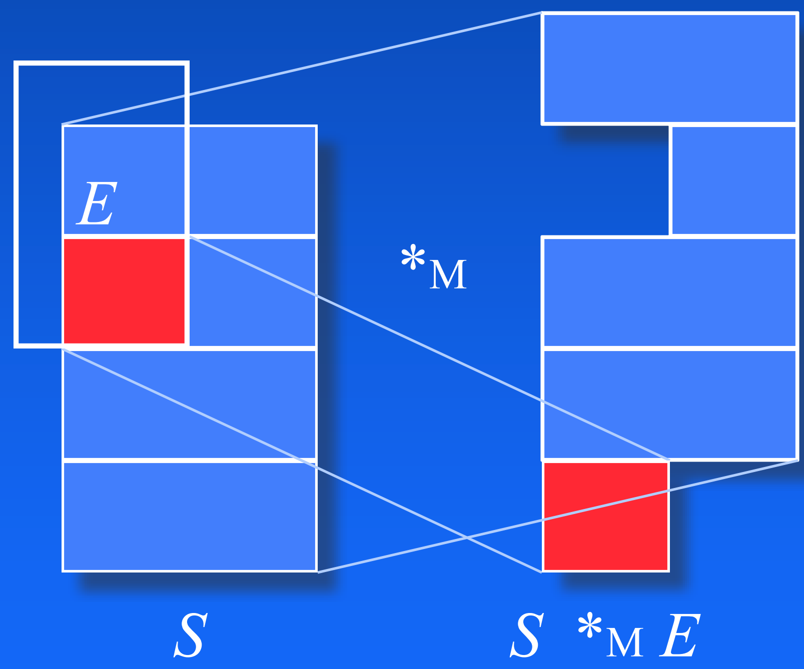
- Drop the lowest possibilities consistent with the data to the bottom and raise everything else up one notch



Minimal or "Natural" Updating $*_M$

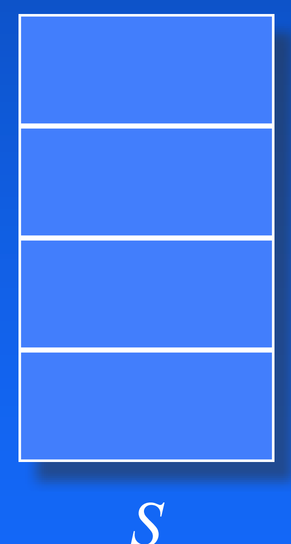
Spohn 88, Boutilier 93

- Drop the lowest possibilities consistent with the data to the bottom and raise everything else up one notch
- inductive leaps
- No epistemic hell



The Flush-to-α Method $*_{F,\alpha}$

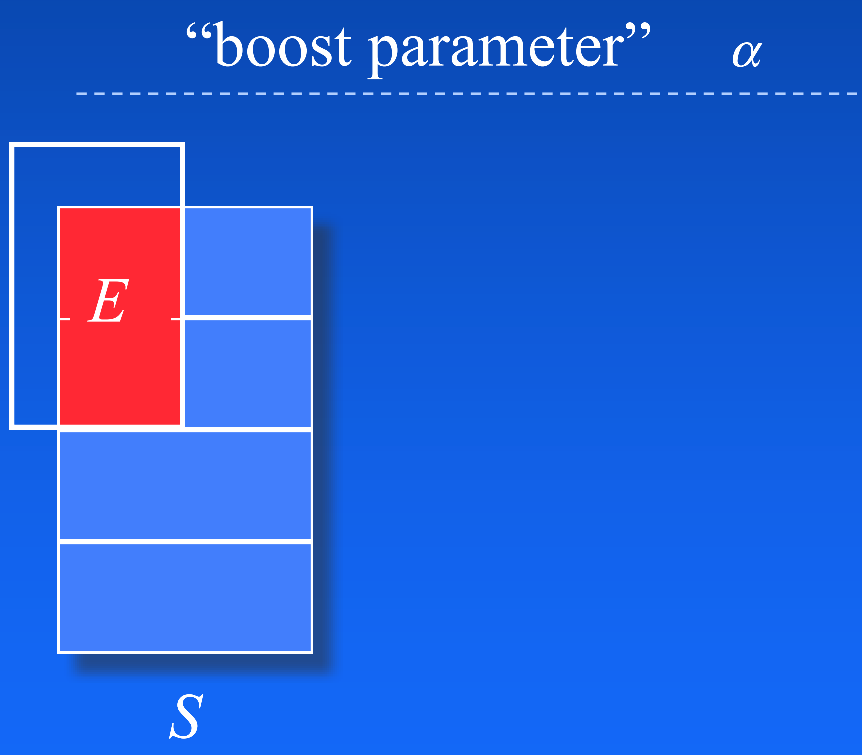
Goldszmidt and Pearl 94



The Flush-to-α Method $*_{F,\alpha}$

Goldszmidt and Pearl 94

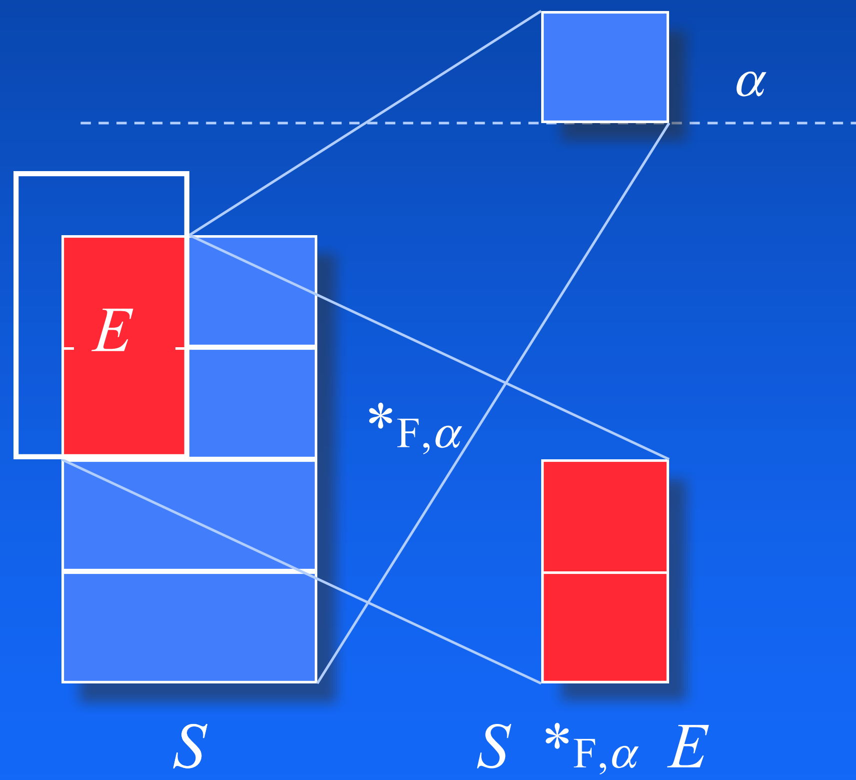
- Send non-E worlds to α and drop E -worlds rigidly to the bottom



The Flush-to-α Method $*_{F,\alpha}$

Goldszmidt and Pearl 94

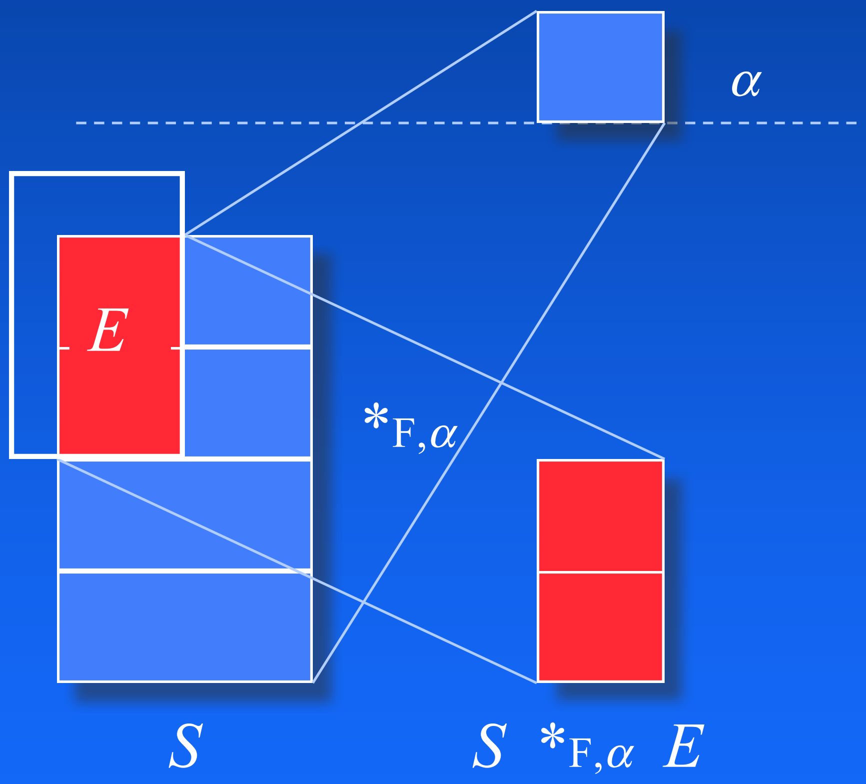
- Send non-E worlds to α and drop E -worlds rigidly to the bottom



The Flush-to-α Method $*_{F,\alpha}$

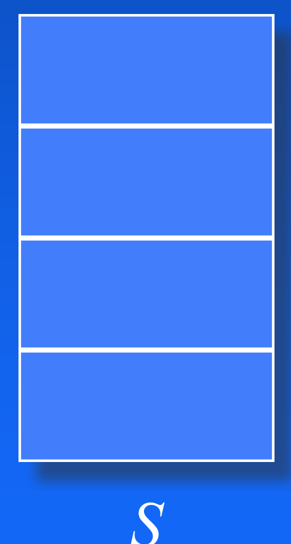
Goldszmidt and Pearl 94

- Send non-E worlds to α and drop E -worlds rigidly to the bottom
- Perfect memory on sequentially consistent data *if* α is high enough
- Inductive leaps
- No epistemic hell



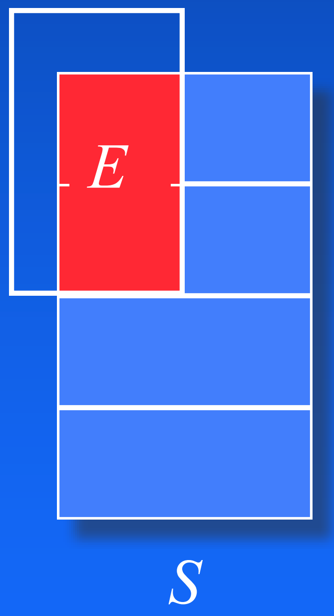
Ordinal Jeffrey Conditioning $*_{J,\alpha}$

Spohn 88



Ordinal Jeffrey Conditioning $*_{J,\alpha}$

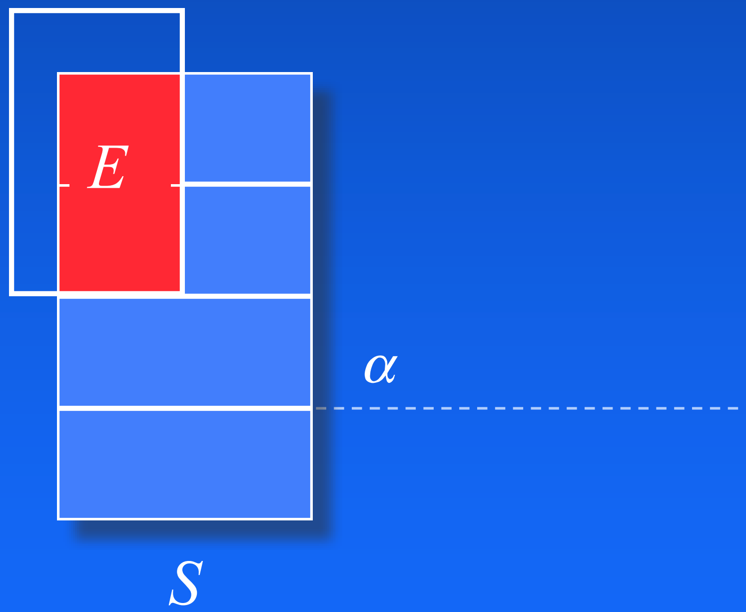
Spohn 88



Ordinal Jeffrey Conditioning $*_{J,\alpha}$

Spohn 88

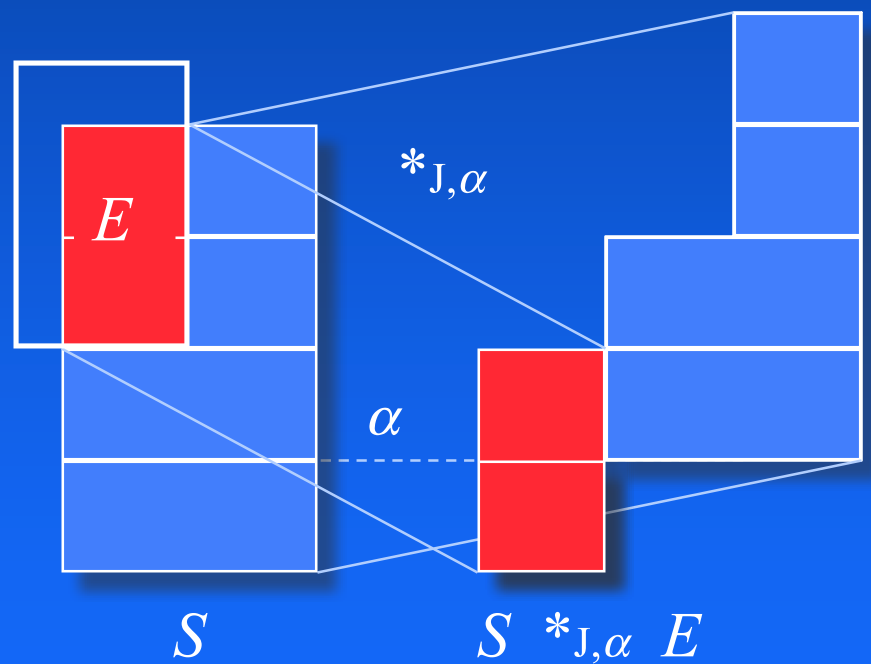
- Drop E worlds to the bottom. Drop non-E worlds to the bottom and then jack them up to level α



Ordinal Jeffrey Conditioning $*_{J,\alpha}$

Spohn 88

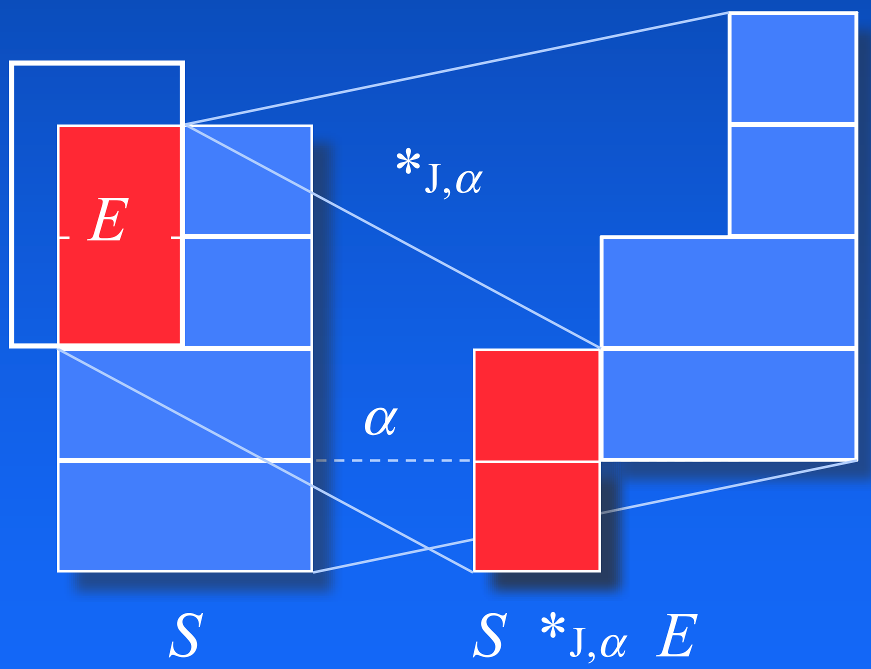
- Drop E worlds to the bottom. Drop non-E worlds to the bottom and then jack them up to level α



Ordinal Jeffrey Conditioning $*_{J,\alpha}$

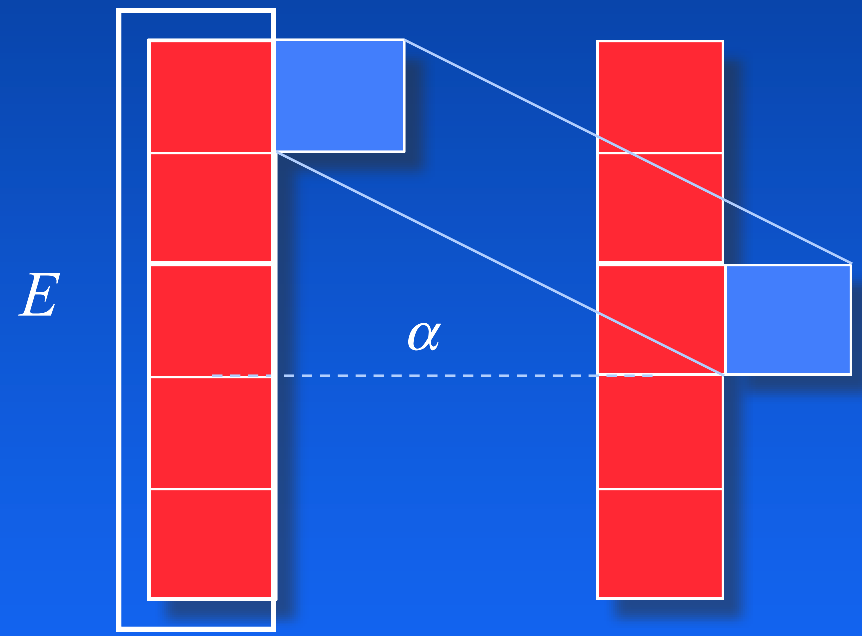
Spohn 88

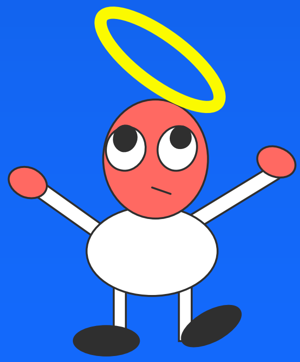
- Drop E worlds to the bottom. Drop non-E worlds to the bottom and then jack them up to level α
- Perfect memory on consistent sequences if α is large enough
- No epistemic hell
- But...



Empirical Backsliding

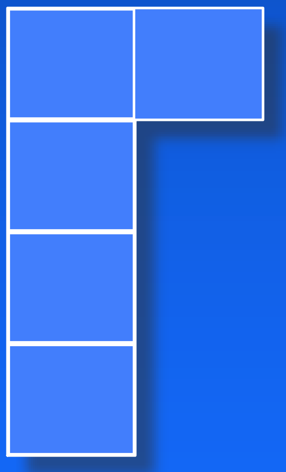
- Ordinal Jeffrey conditioning can increase the plausibility of a refuted possibility





The Ratchet Method $*_{R,\alpha}$

Darwiche and Pearl 97

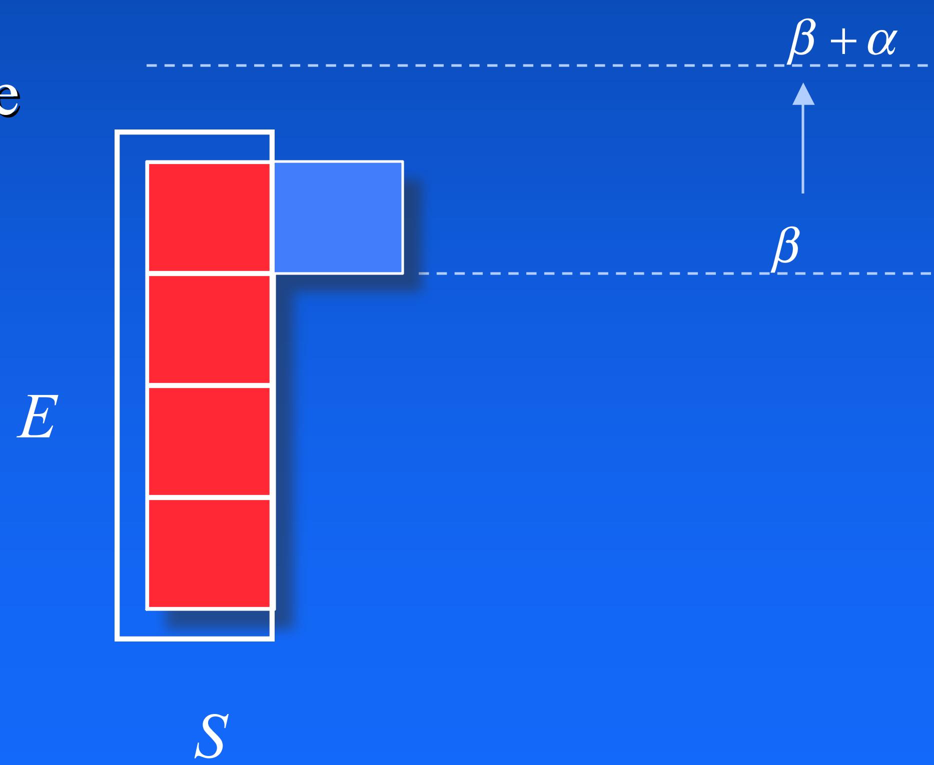


S

The Ratchet Method $*_{R,\alpha}$

Darwiche and Pearl 97

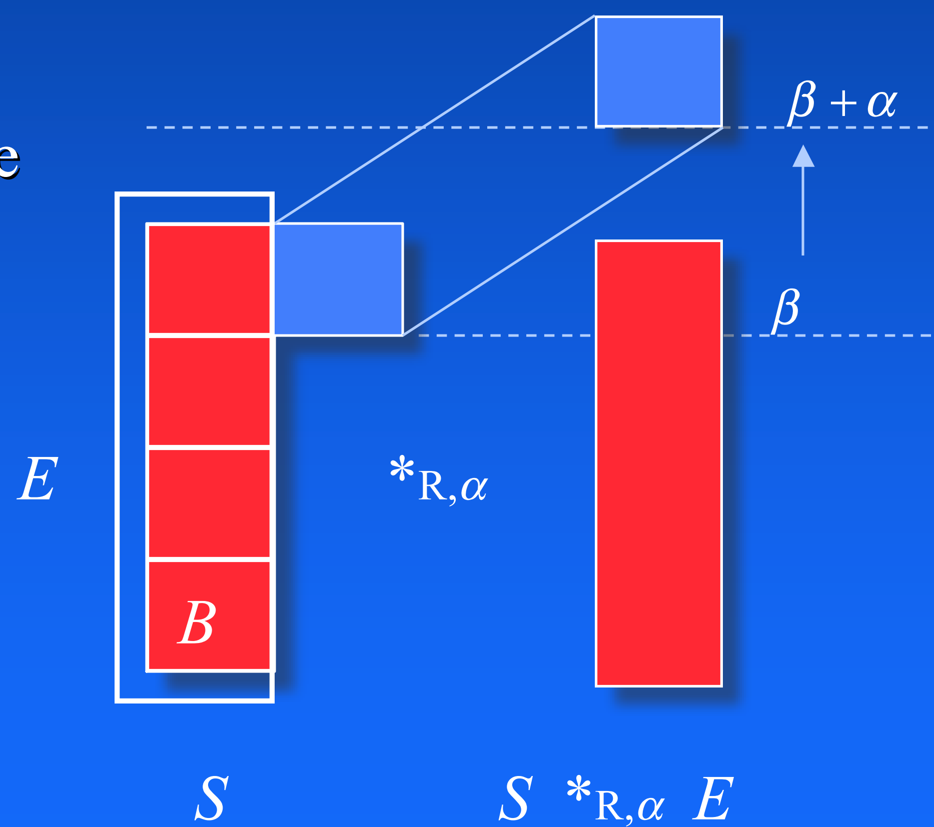
- Like ordinal Jeffrey conditioning except refuted possibilities move up by α from their current positions



The Ratchet Method $*_{R,\alpha}$

Darwiche and Pearl 97

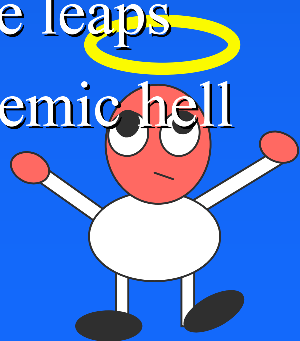
- Like ordinal Jeffrey conditioning except refuted possibilities move up by α from their current positions

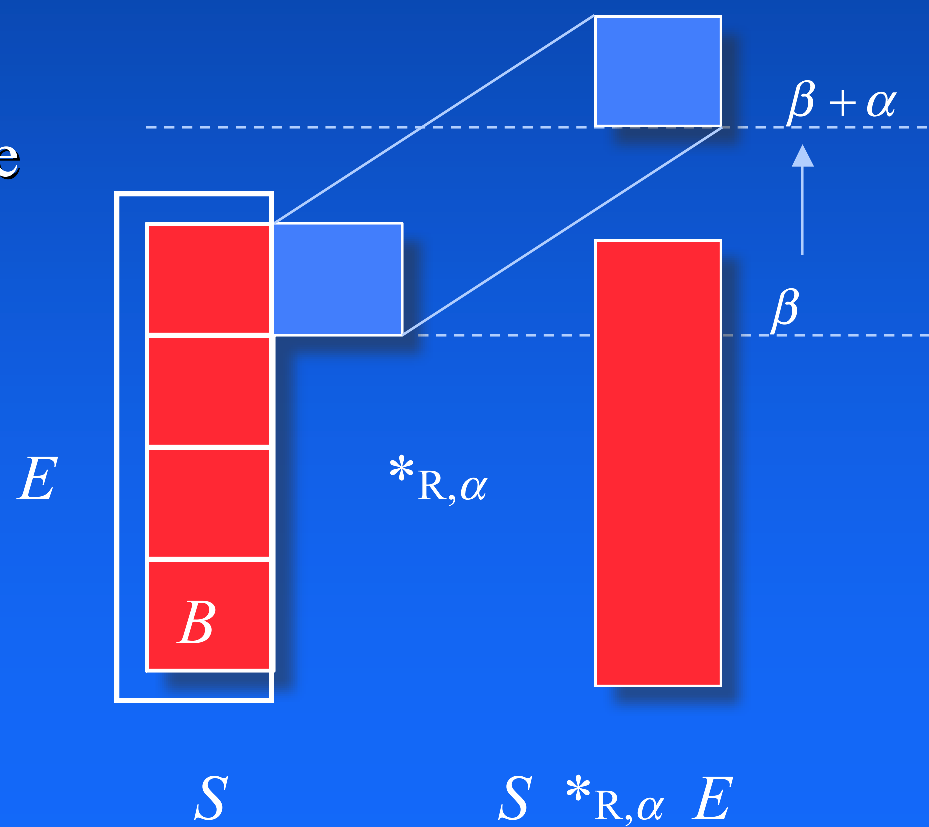


The Ratchet Method $*_{R,\alpha}$

Darwiche and Pearl 97

- Like ordinal Jeffrey conditioning except refuted possibilities move up by α from their current positions
- Perfect memory if α is large enough
- Inductive leaps
- No epistemic hell





Part II

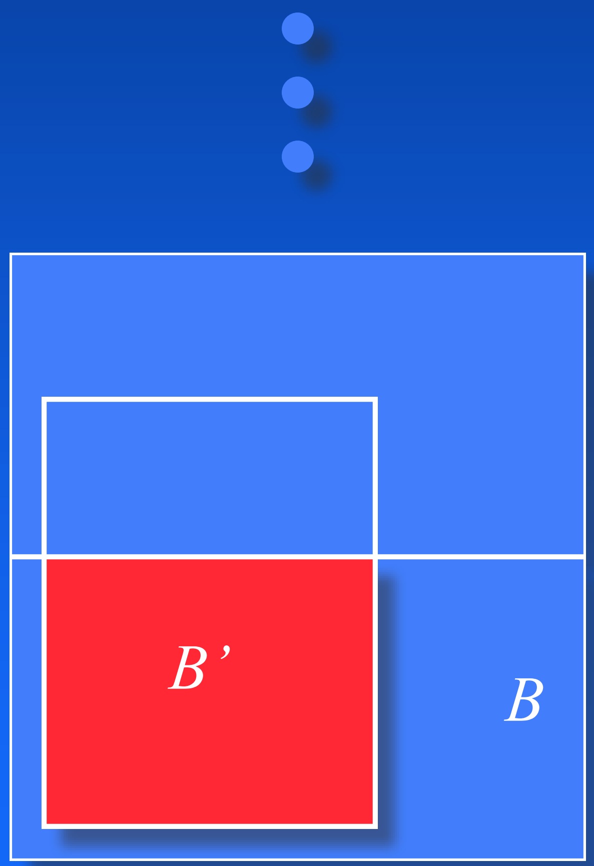
Properties of the Methods

Timidity and Stubbornness

- Timidity: *no inductive leaps without refutation.*
- Stubbornness: *no retractions without refutation*

- Examples: *all the above*

- **Nutty!**

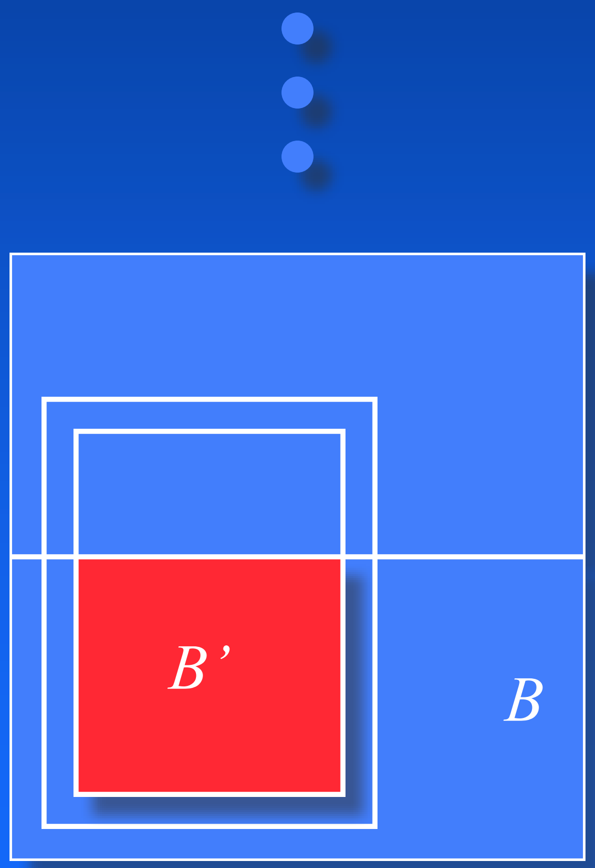


Timidity and Stubbornness

- Timidity: *no inductive leaps without refutation.*
- Stubbornness: *no retractions without refutation*

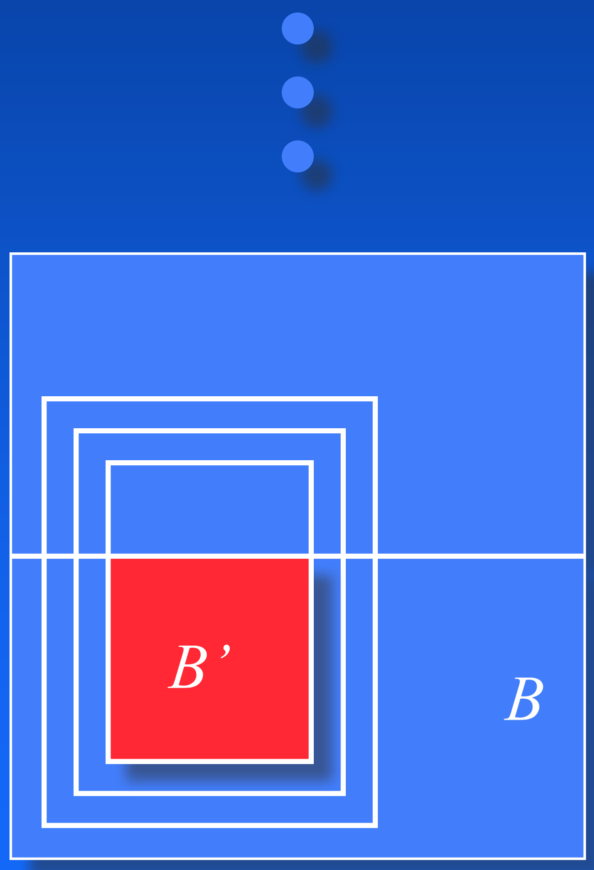
- Examples: *all the above*

- **Nutty!**



Timidity and Stubbornness

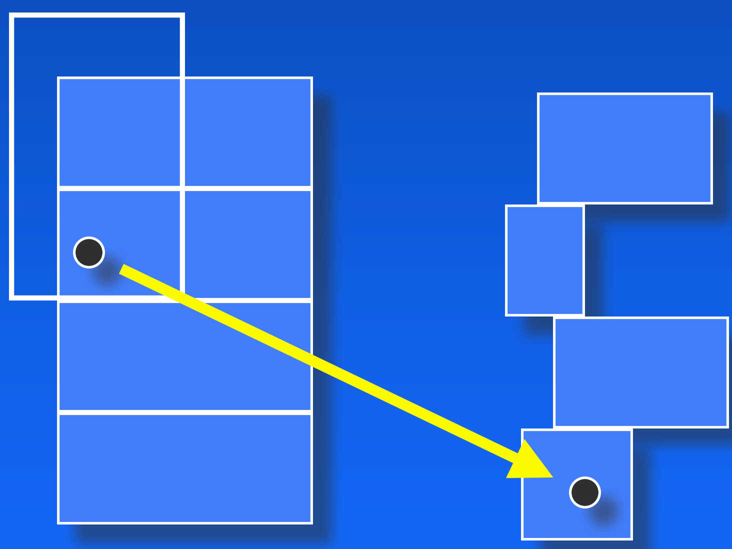
- Timidity: *no inductive leaps without refutation.*
- Stubbornness: *no retractions without refutation*
- Examples: *all the above*
- **Nutty!**



Local Consistency

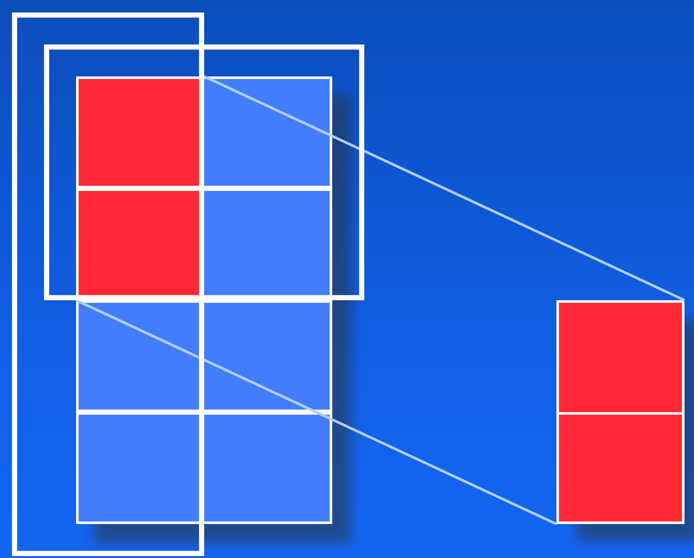
- Local consistency: *new belief must be consistent with the current consistent datum*

- Examples: *all the above*



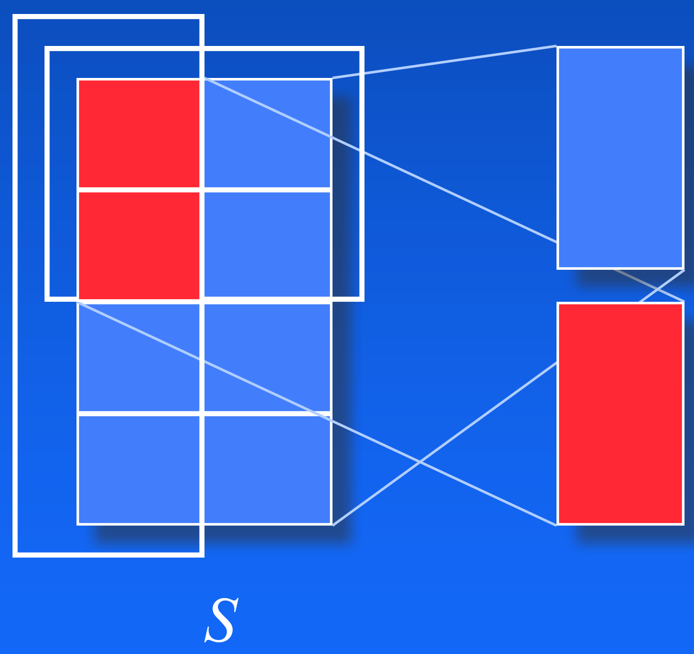
Positive Order-invariance

- Positive order-invariance: *preserve original ranking inside conjunction of data*
- Examples:
  - $*_{C}$, $*_{L}$, $*_{R,\alpha}$, $*_{J,\alpha}$.



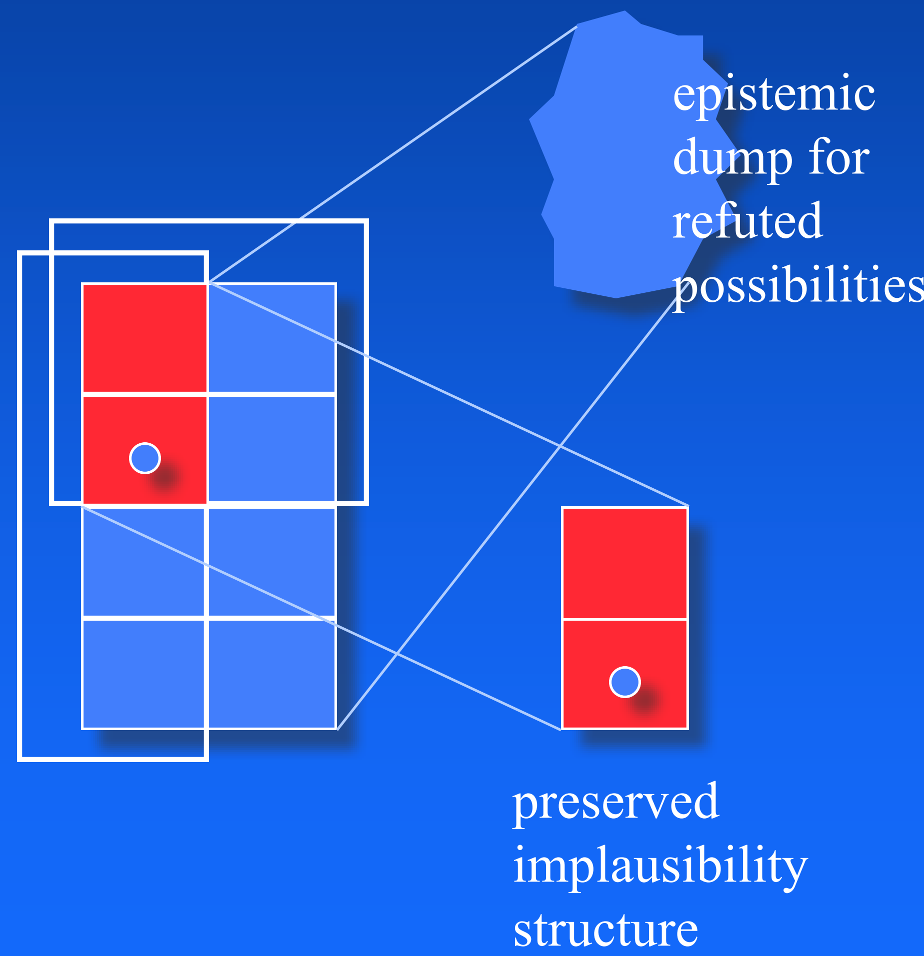
Data-Precedence

- Data-precedence: *Each world satisfying all the data is placed above each world failing to satisfy some datum.*
- Examples:
  - $*_{C}$, $*_{L}$
  - $*_{R,\, \alpha}$, $*_{J,\, \alpha}$, if α is above S .



Enumerate and Test

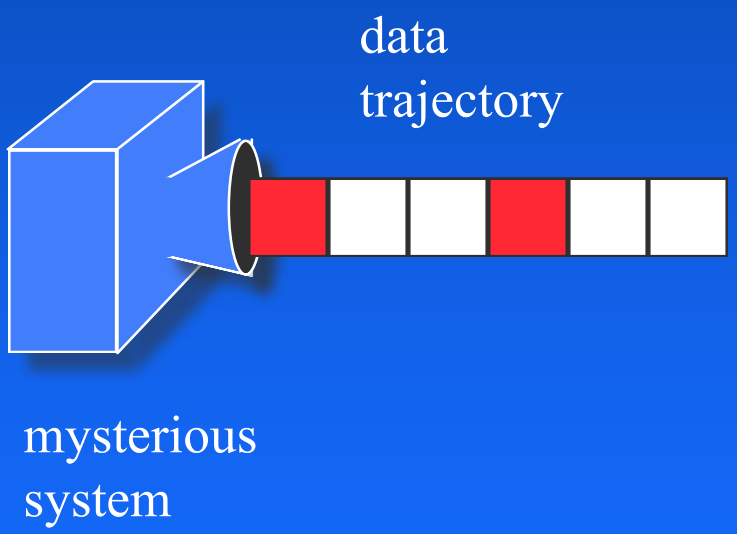
- Enumerate-and-test:
  - *locally consistent,*
  - *positively invariant*
  - *data-precedent*
- Examples:
  - $*_C$, $*_L$
  - $*_{R, \alpha}$, $*_{J, \alpha}$, if α is above S .



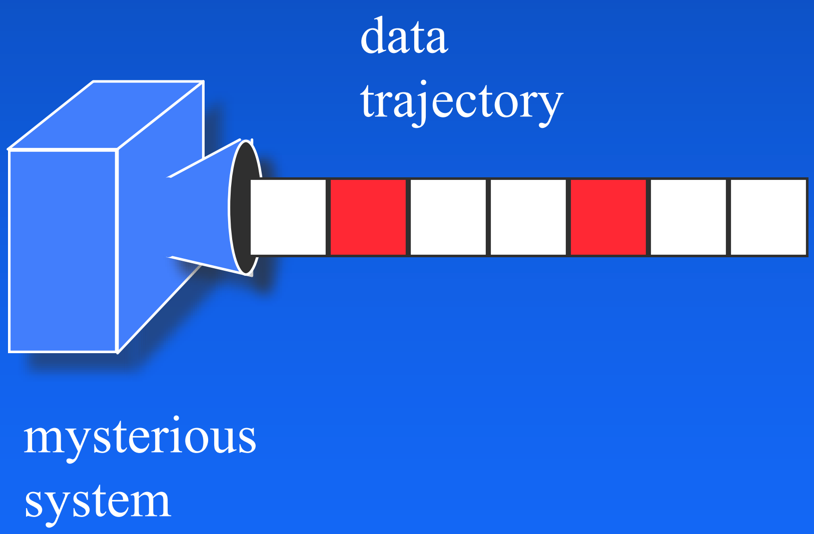
Part III

Belief Revision as Learning

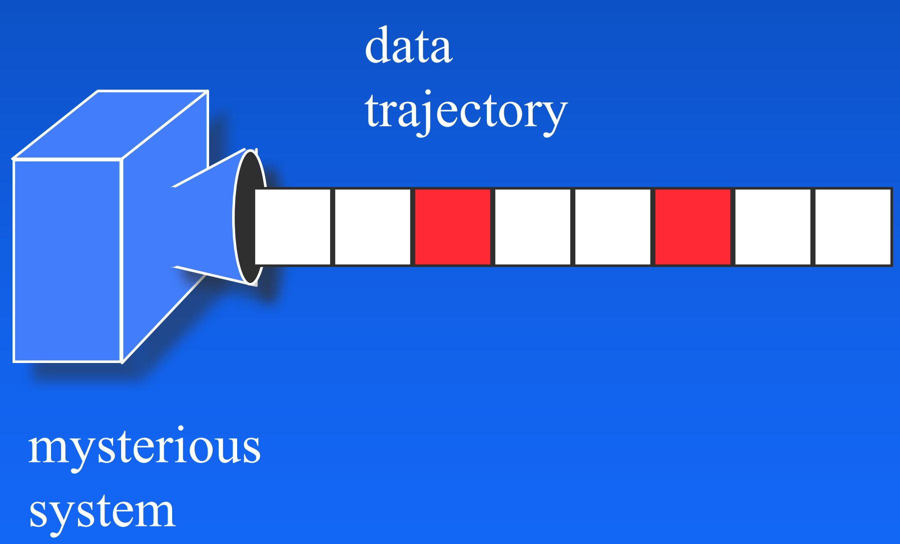
A Very Simple Learning Paradigm



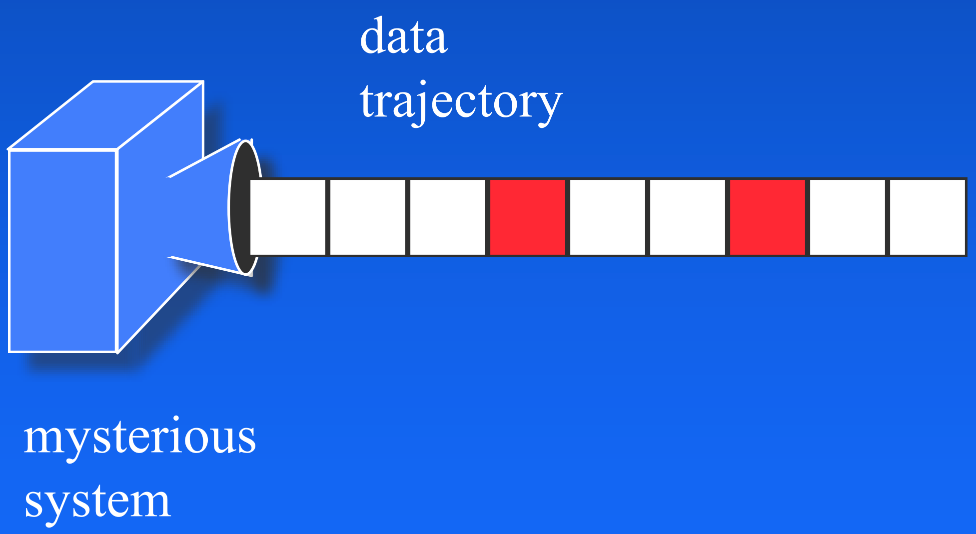
A Very Simple Learning Paradigm



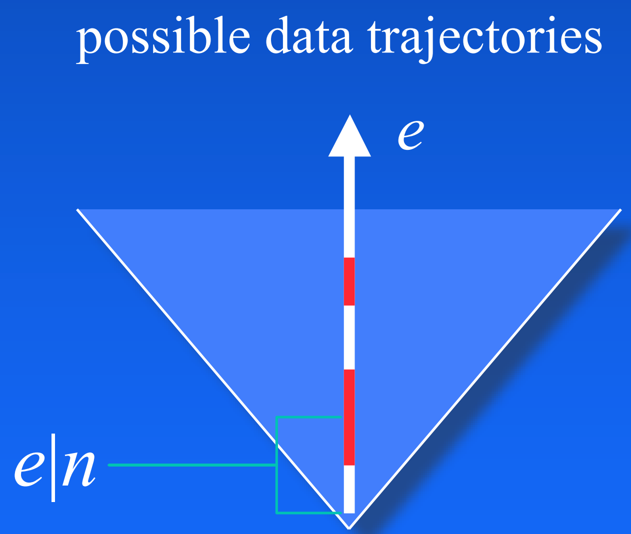
A Very Simple Learning Paradigm



A Very Simple Learning Paradigm



Possible Outcome Trajectories



Finding the Truth

$(*, S_0)$ identifies $e \Leftrightarrow$

for all but finitely many n,

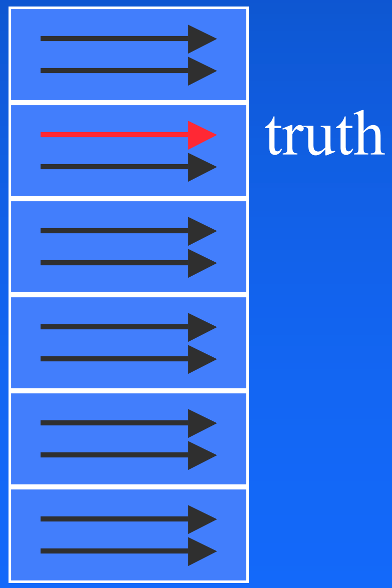
$$b(S_0 * ([0, e(0)], \ldots, [n, e(n)])) = \{e\}$$

Finding the Truth

$(*, S_0)$ identifies $e \Leftrightarrow$

for all but finitely many n,

$$b(S_0 * ([0, e(0)], \ldots, [n, e(n)])) = \{e\}$$

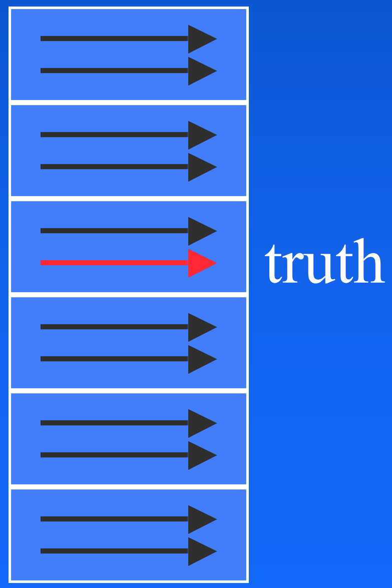


Finding the Truth

$(*, S_0)$ identifies e $\Leftrightarrow$

for all but finitely many n,

$$b(S_0 * ([0, e(0)], \ldots , [n, e(n)])) = \{e\}$$

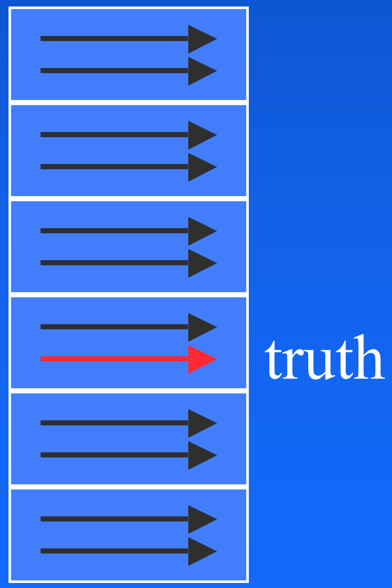


Finding the Truth

$(*, S_0)$ identifies $e \Leftrightarrow$

for all but finitely many n,

$b(S_0 * ([0, e(0)], \ldots , [n, e(n)])) = \{e\}$

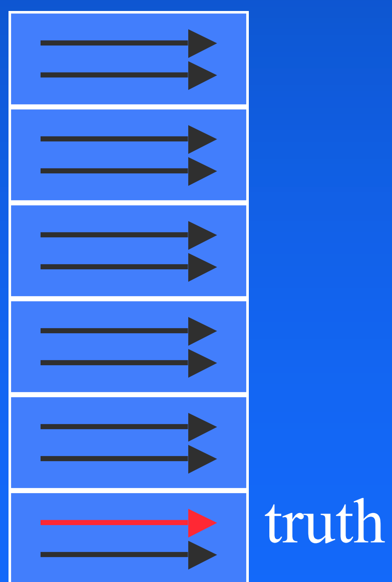


Finding the Truth

$(*, S_0)$ identifies $e \Leftrightarrow$

for all but finitely many n,

$b(S_0 * ([0, e(0)], \ldots, [n, e(n)]) = \{e\}$

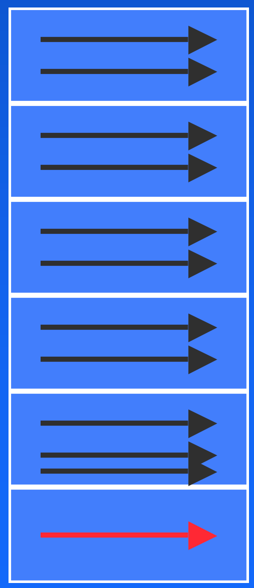


Finding the Truth

$(*, S_0)$ identifies $e \Leftrightarrow$

for all but finitely many n,

$$b(S_0 * ([0, e(0)], \ldots, [n, e(n)]) = \{e\}$$



completely true belief

Reliability is No Accident

- Let K be a range of possible outcome trajectories
- $(*, S_0)$ *identifies* $K \Leftrightarrow (*, S_0)$ identifies each e in K.

- Fact: K is identifiable $\Leftrightarrow$ K is countable.

Completeness

- * is complete $\Leftrightarrow$
- for each identifiable K
- there is an S_0 such that,
- K is identifiable by $(*, S_0)$.

- Else * is restrictive.

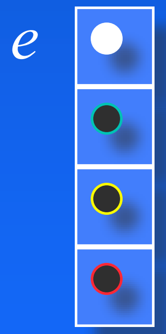
Completeness

Proposition: If * enumerates and tests, * is complete.

Completeness

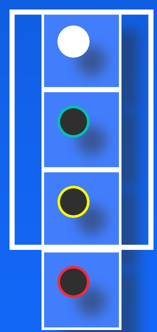
Proposition: If * enumerates and tests, * is complete.

- Enumerate K
- Choose arbitrary e in K



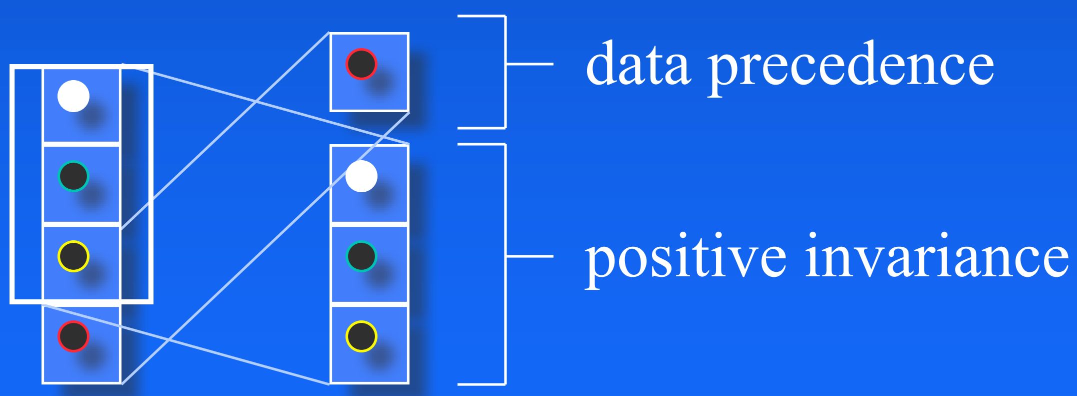
Completeness

Proposition: If * enumerates and tests, * is complete.



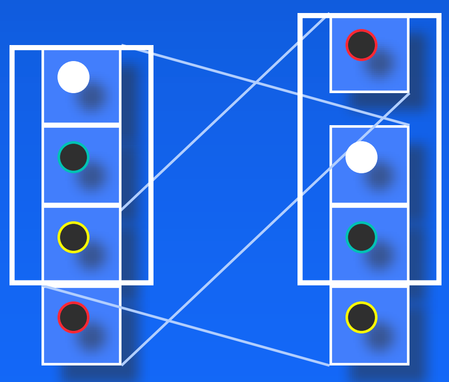
Completeness

Proposition: If * enumerates and tests, * is complete.



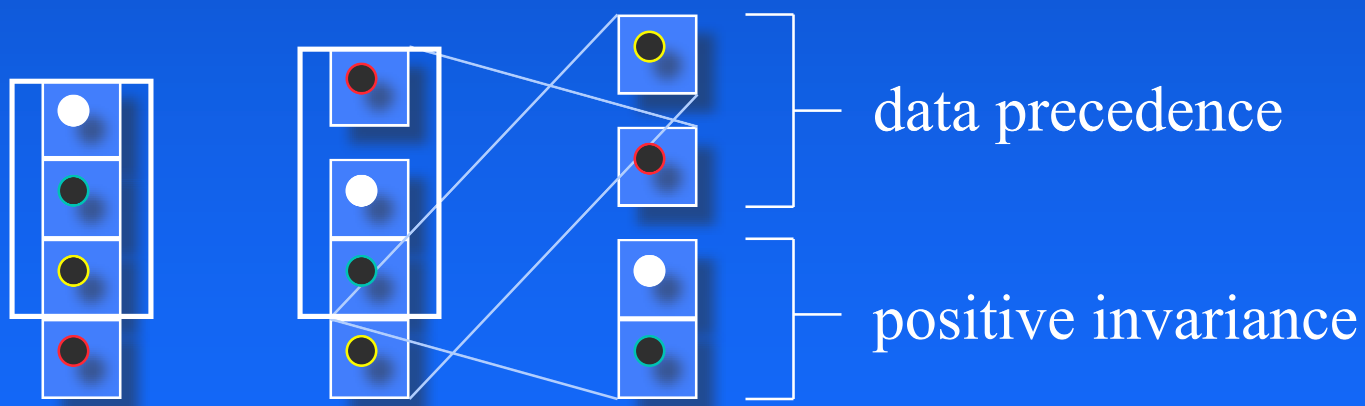
Completeness

Proposition: If * enumerates and tests, * is complete.



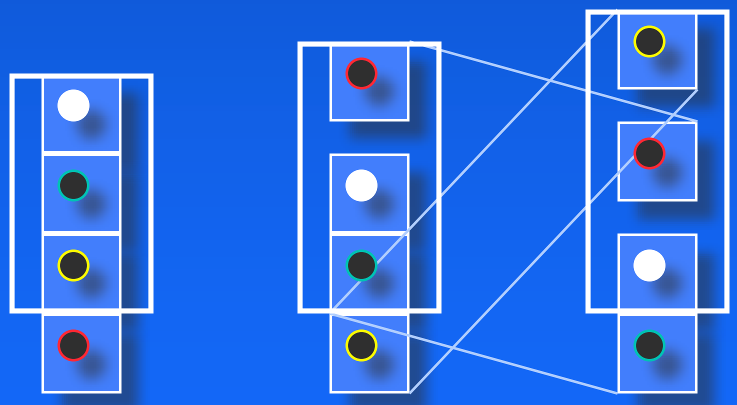
Completeness

Proposition: If * enumerates and tests, * is complete.



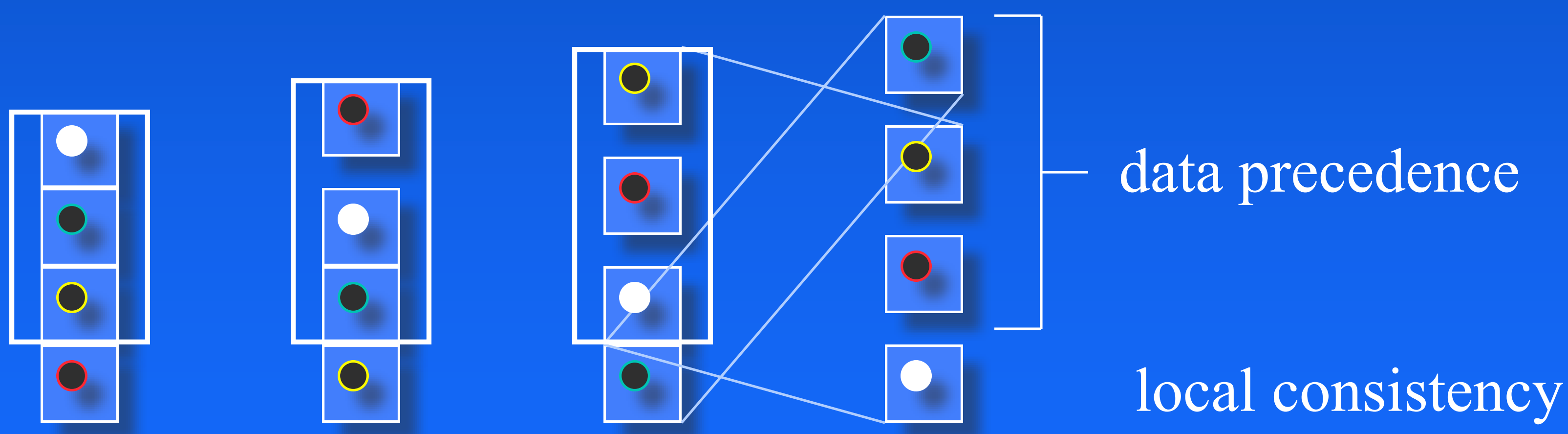
Completeness

Proposition: If * enumerates and tests, * is complete.



Completeness

Proposition: If * enumerates and tests, * is complete.

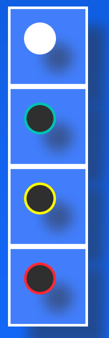


convergence

Amnesia

Without data precedence, memory can fail

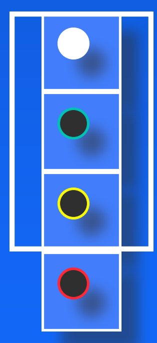
Same example, using $*_{J,1}$.



Amnesia

Without data precedence, memory can fail

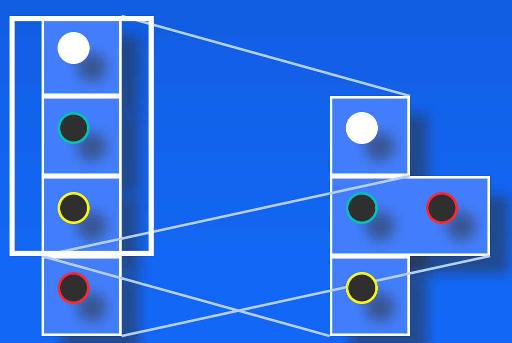
Same example, using $*_{J,1}$.



Amnesia

Without data precedence, memory can fail

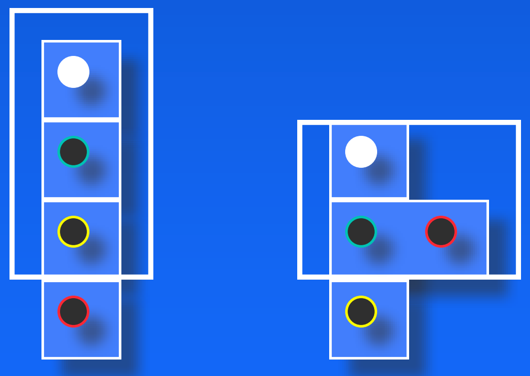
Same example, using $*_{J,1}$.



Amnesia

Without data precedence, memory can fail

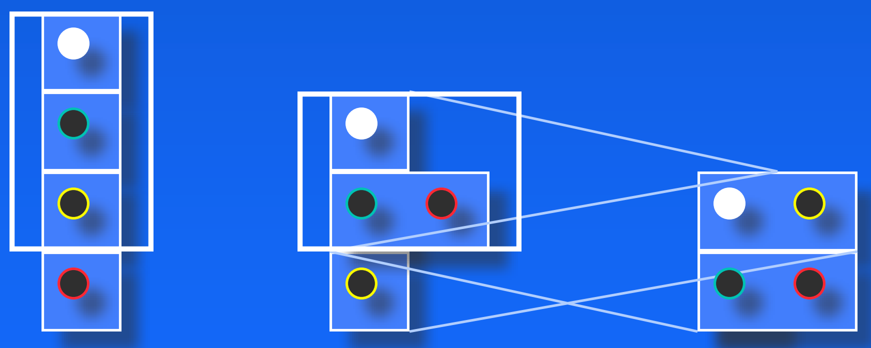
Same example, using $*_{J,1}$.



Amnesia

Without data precedence, memory can fail

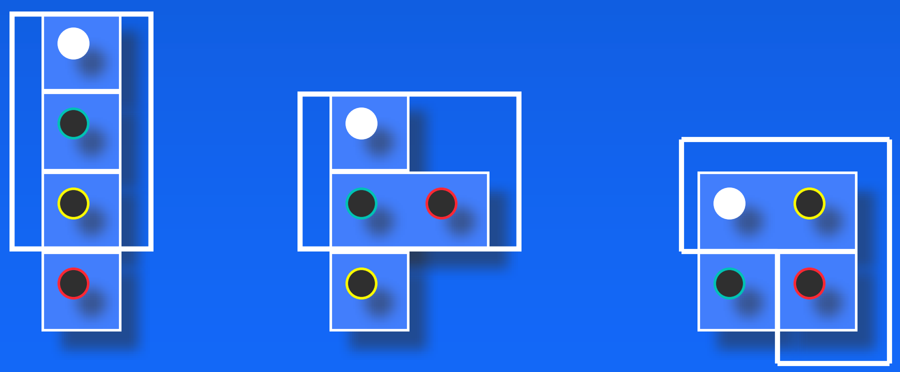
Same example, using $*_{J,1}$.



Amnesia

Without data precedence, memory can fail

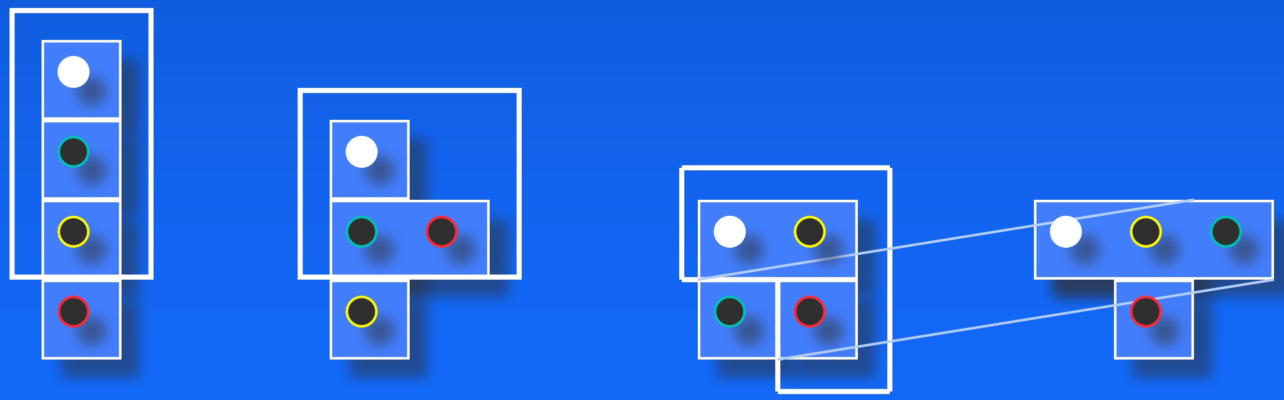
Same example, using $*_{J,1}$.



Amnesia

Without data precedence, memory can fail

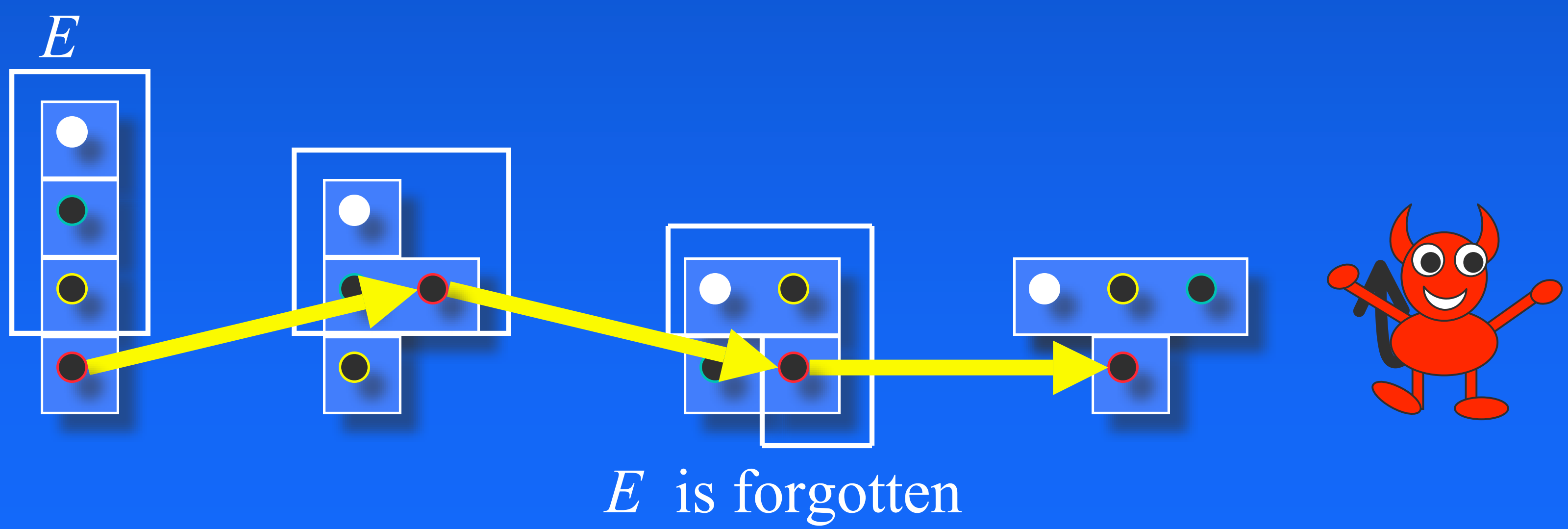
Same example, using $*_{J,1}$.



Amnesia

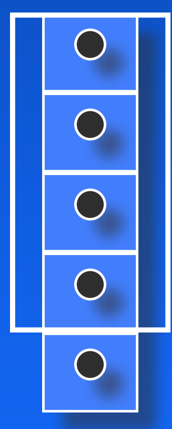
Without data precedence, memory can fail

Same example, using $*_{J,1}$.



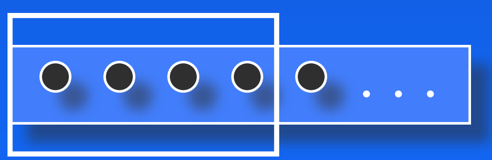
Duality

conjectures and refutations



predicts
may forget

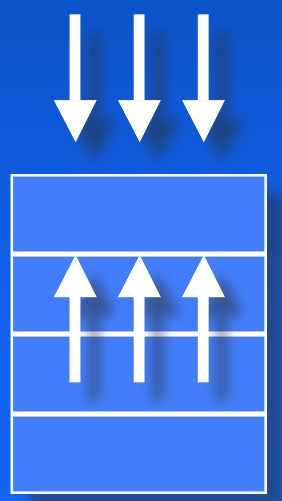
tabula rasa



remembers
doesn’t predict

“Rationally” Imposed Tension

compression for memory



Can both be accommodated?

rarefaction for inductive leaps

Inductive Amnesia

compression for memory

rarefaction for inductive leaps

Restrictiveness:
No possible initial state resolves the pressure

Question

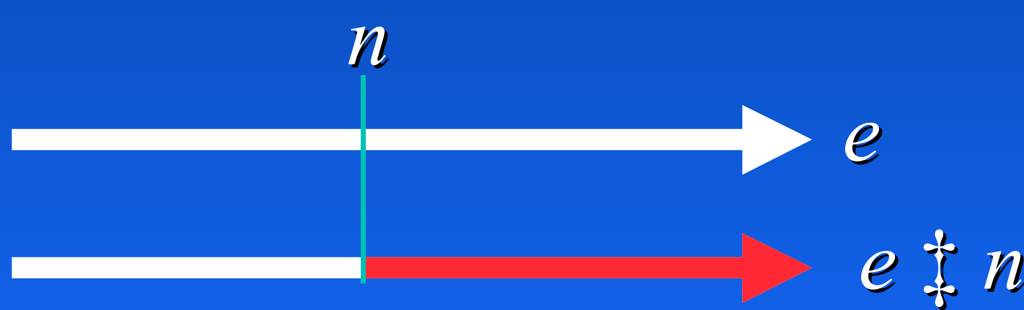
- Which methods are guilty?
- Are some worse than others?

Part IV:

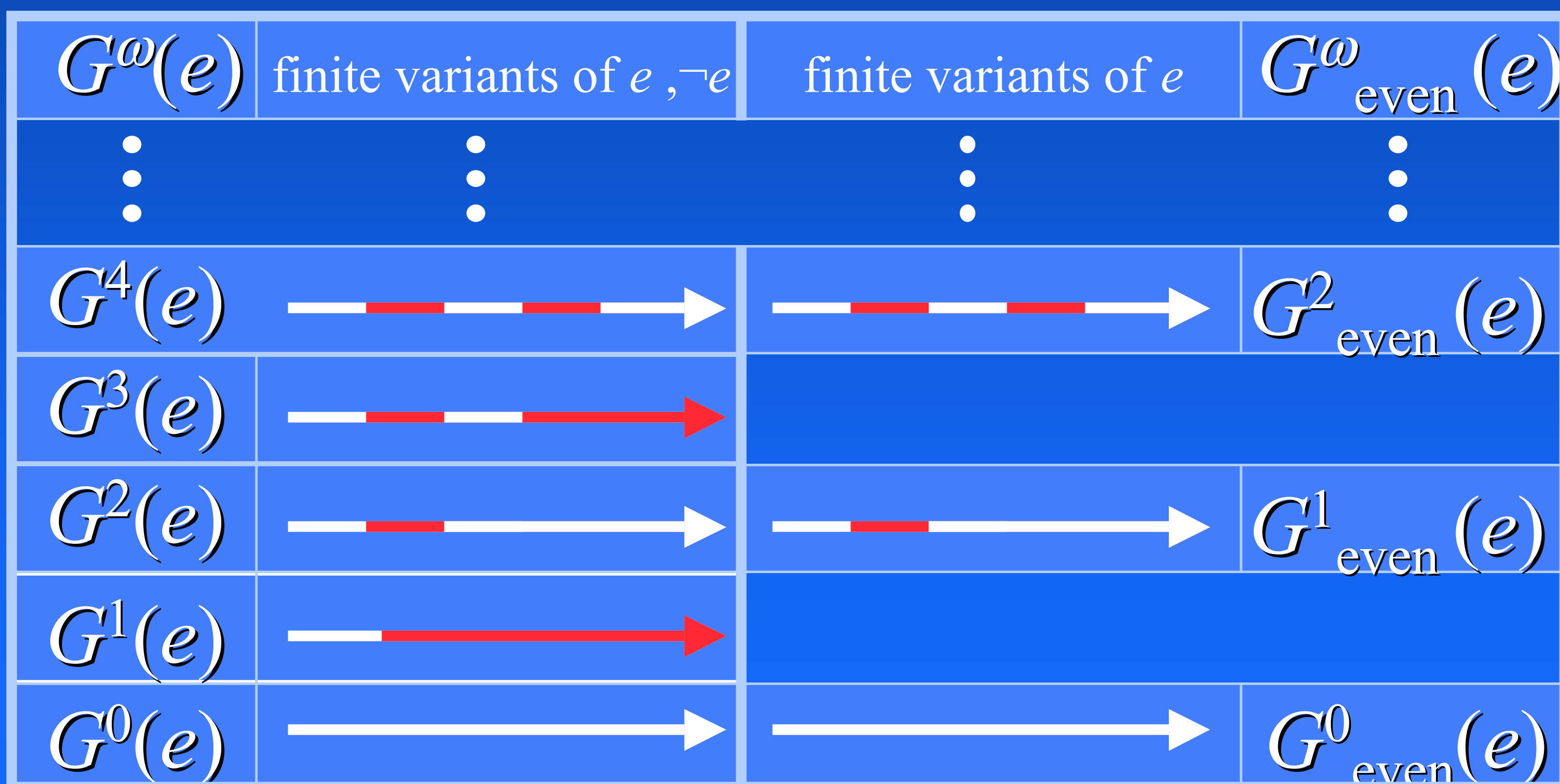
The Goodman Hierarchy

The Grue Operation

Nelson Goodman



Grue Complexity Hierarchy



Classification: even grues

	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}{}_{\text{even}}(e)$	*no*	$\alpha = \omega$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
⋮						
$G^{n}{}_{\text{even}}(e)$	*no*	$\alpha = n+1$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
⋮						
$G^{2}{}_{\text{even}}(e)$	*no*	$\alpha = 3$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$G^{1}{}_{\text{even}}(e)$	*no*	$\alpha = 2$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$G^{0}{}_{\text{even}}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

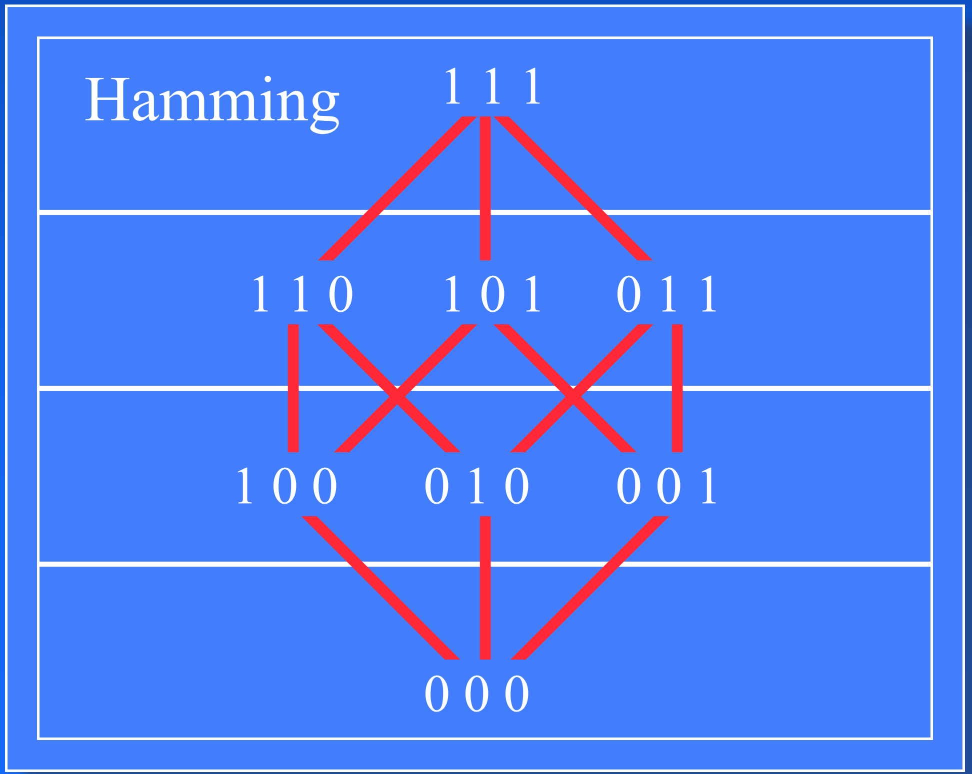
Classification: even grues

	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}{}_{\text{even}}(e)$	*no*	$\alpha = \omega$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
⋮			⋮			
$G^{n}{}_{\text{even}}(e)$	*no*	$\alpha = n+1$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
⋮			⋮			
$G^{2}{}_{\text{even}}(e)$	*no*	$\alpha = 3$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$G^{1}{}_{\text{even}}(e)$	*no*	$\alpha = 2$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$G^{0}{}_{\text{even}}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

Hamming Algebra

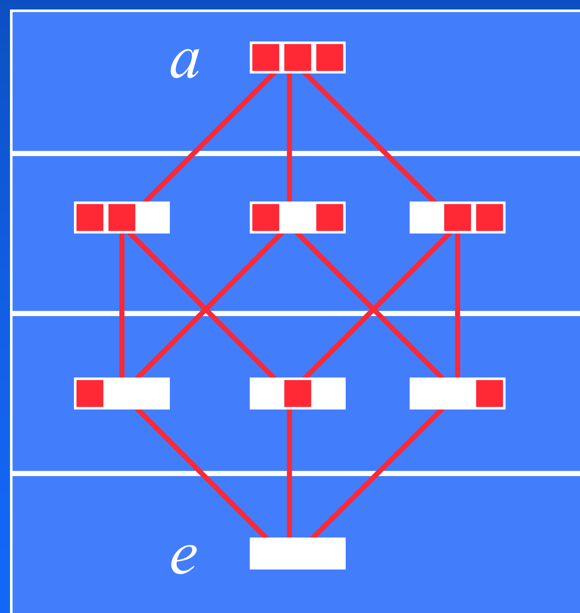
- $a \leq_H b \ mod \ e \Leftrightarrow$

  a differs from *e* only where *b* does.



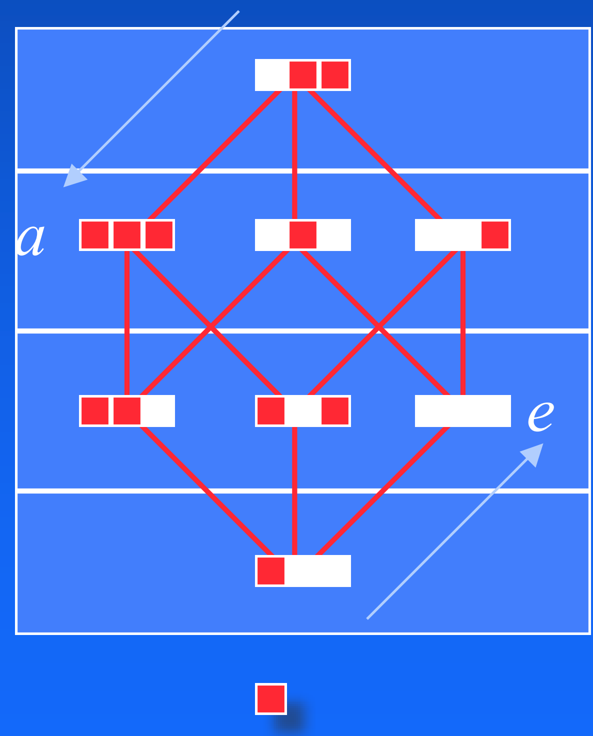
$*_{R,1}$,$*_{J,1}$ can identify $G^{\omega}_{\text{even}}(e)$

Example



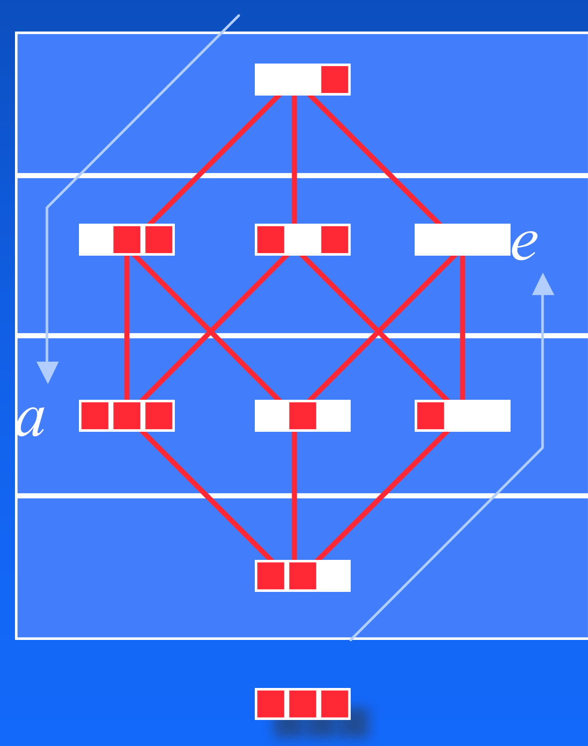
Learning as rigid hypercube rotation

$*_{R,1}$,$*_{J,1}$ can identify $G^{\omega}{}_{\text{even}}(e)$



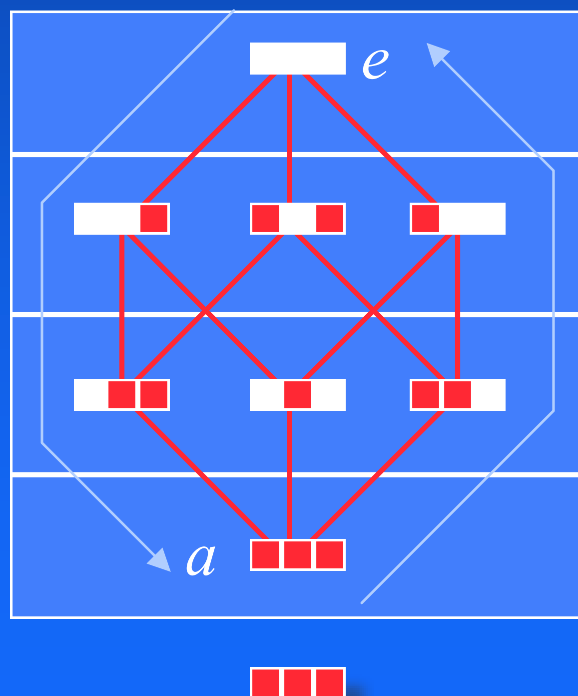
Learning as rigid hypercube rotation

$*_{R,1}\,,*_{J,1}$ can identify $G^{\omega}{}_{even}(e)$



Learning as rigid hypercube rotation

$*_{R,1}$,$*_{J,1}$ can identify $G^{\omega}{}_{even}(e)$



convergence

Learning as rigid hypercube rotation

Classification: even grues

	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}{}_{\text{even}}(e)$	*no*	$\alpha = \omega$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
⋮			⋮			
$G^{n}{}_{\text{even}}(e)$	*no*	$\alpha = n+1$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
⋮			⋮			
$G^{2}{}_{\text{even}}(e)$	*no*	$\alpha = 3$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$G^{1}{}_{\text{even}}(e)$	*no*	$\alpha = 2$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$G^{0}{}_{\text{even}}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

Classification: arbitrary grues

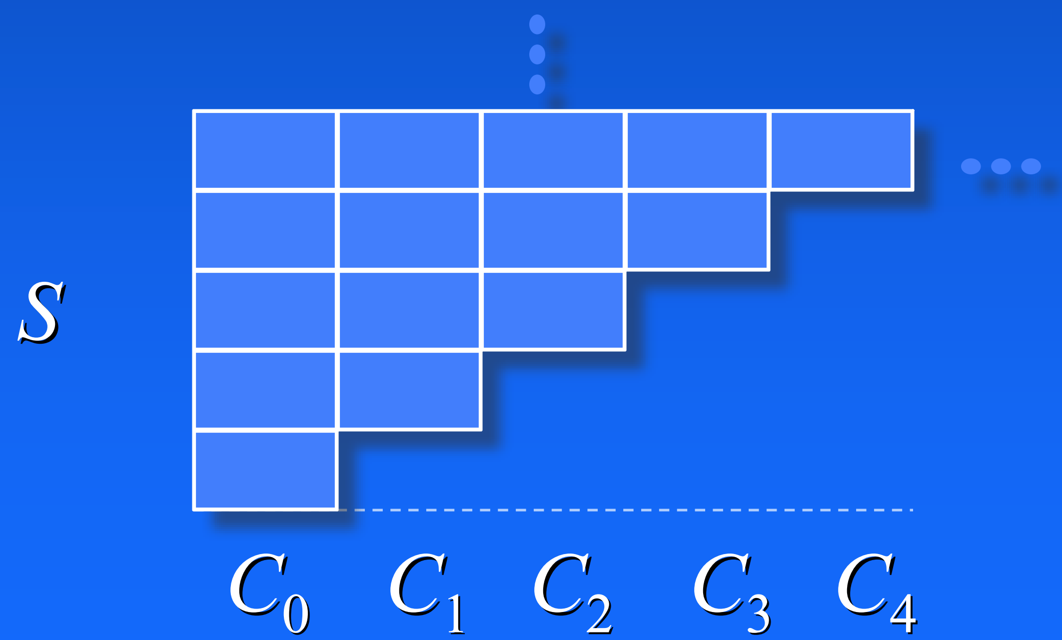
	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}(e)$	*no*	$\alpha = \omega$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$\vdots$						
$G^{3}(e)$	*no*	$\alpha = n+1$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$\vdots$						
$G^{2}(e)$	*no*	$\alpha = 3$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{1}(e)$	*no*	$\alpha = 2$	$\alpha = 2$	$\alpha = 1$	*yes*	*yes*
$G^{0}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

Classification: arbitrary grues

	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}(e)$	*no*	$\alpha = \omega$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$\vdots$			$\vdots$			
$G^{3}(e)$	*no*	$\alpha = n + 1$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$\vdots$			$\vdots$			
$G^{2}(e)$	*no*	$\alpha = 3$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{1}(e)$	*no*	$\alpha = 2$	$\alpha = 2$	$\alpha = 1$	*yes*	*yes*
$G^{0}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

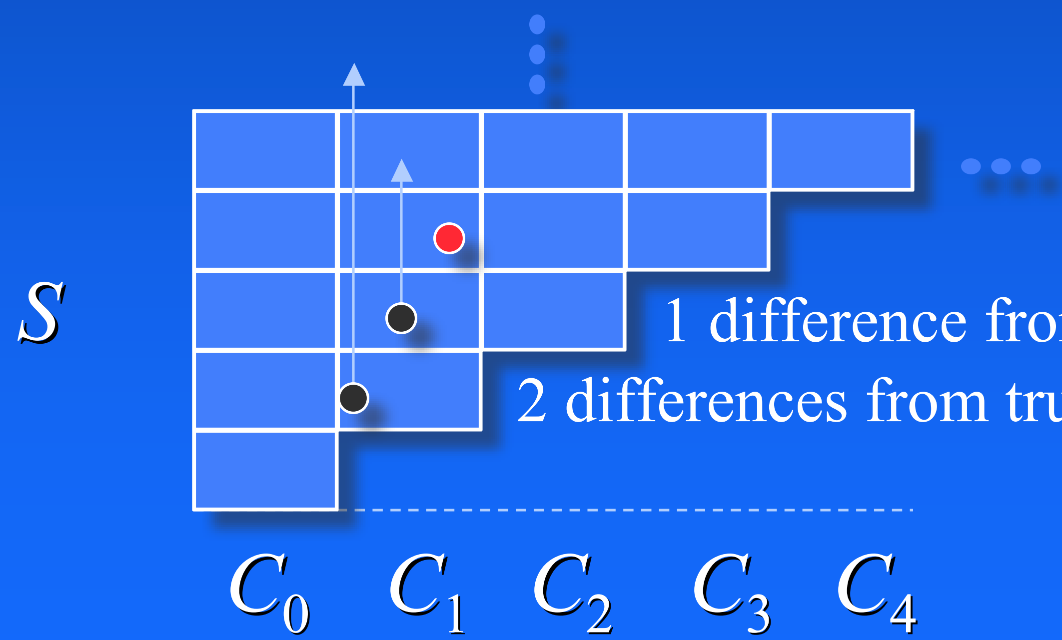
$*_{R,2}$ is Complete

- Impose the Hamming distance ranking on each finite variant class
- Now raise the *n*th Hamming ranking by *n*



$*_{R,2}$ is Complete

- Data streams in the same column just barely make it because they jump by 2 for each difference from the truth



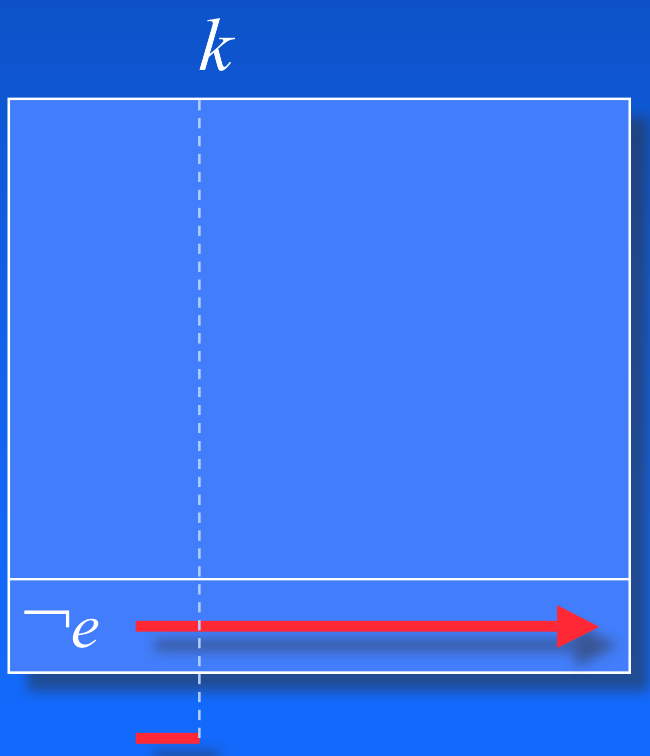
Classification: arbitrary grues

	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}(e)$	*no*	$\alpha = \omega$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{3}(e)$	*no*	$\alpha = n+1$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{2}(e)$	*no*	$\alpha = 3$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{1}(e)$	*no*	$\alpha = 2$	$\alpha = 2$	$\alpha = 1$	*yes*	*yes*
$G^{0}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

Can't use Hamming rank

Wrench In the Works

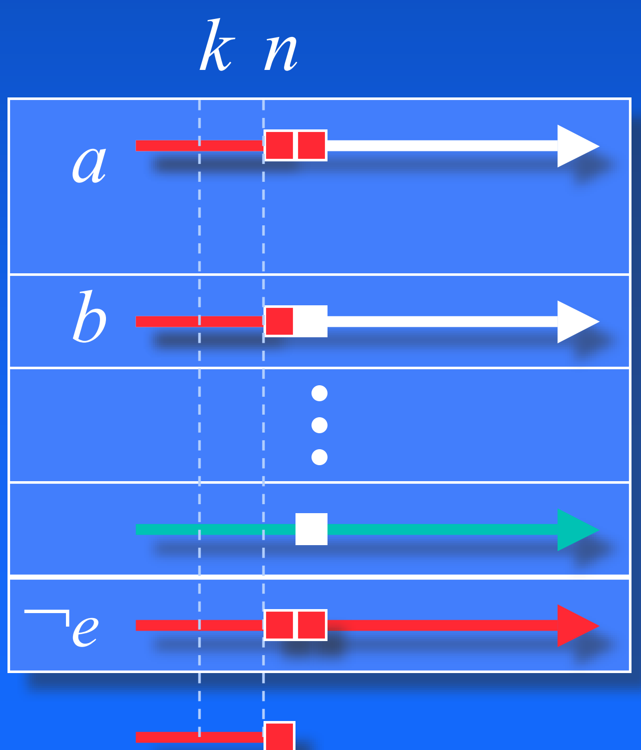
- Suppose $*_{J,2}$ succeeds with Hamming rank.
- Feed $\neg e$ until it is uniquely at the bottom.



By convergent success

Wrench In the Works

- So for some later *n*,



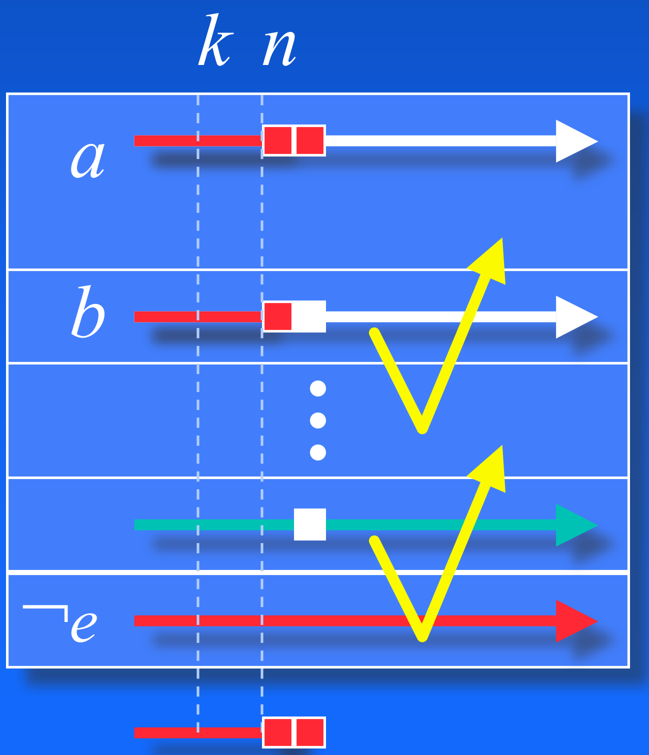
Hamming rank and positive invariance.

If empty, things go even worse!

Still alone since timid and stubborn

Wrench In the Works

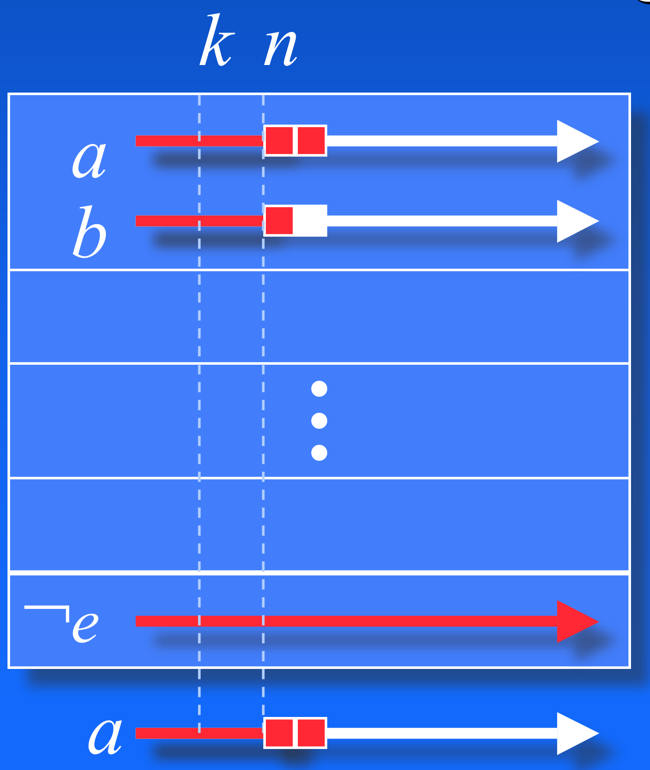
- *b* moves up at most 1 step since ¬*e* is still alone (rule)



Refuted worlds touch bottom and get lifted by at most two.

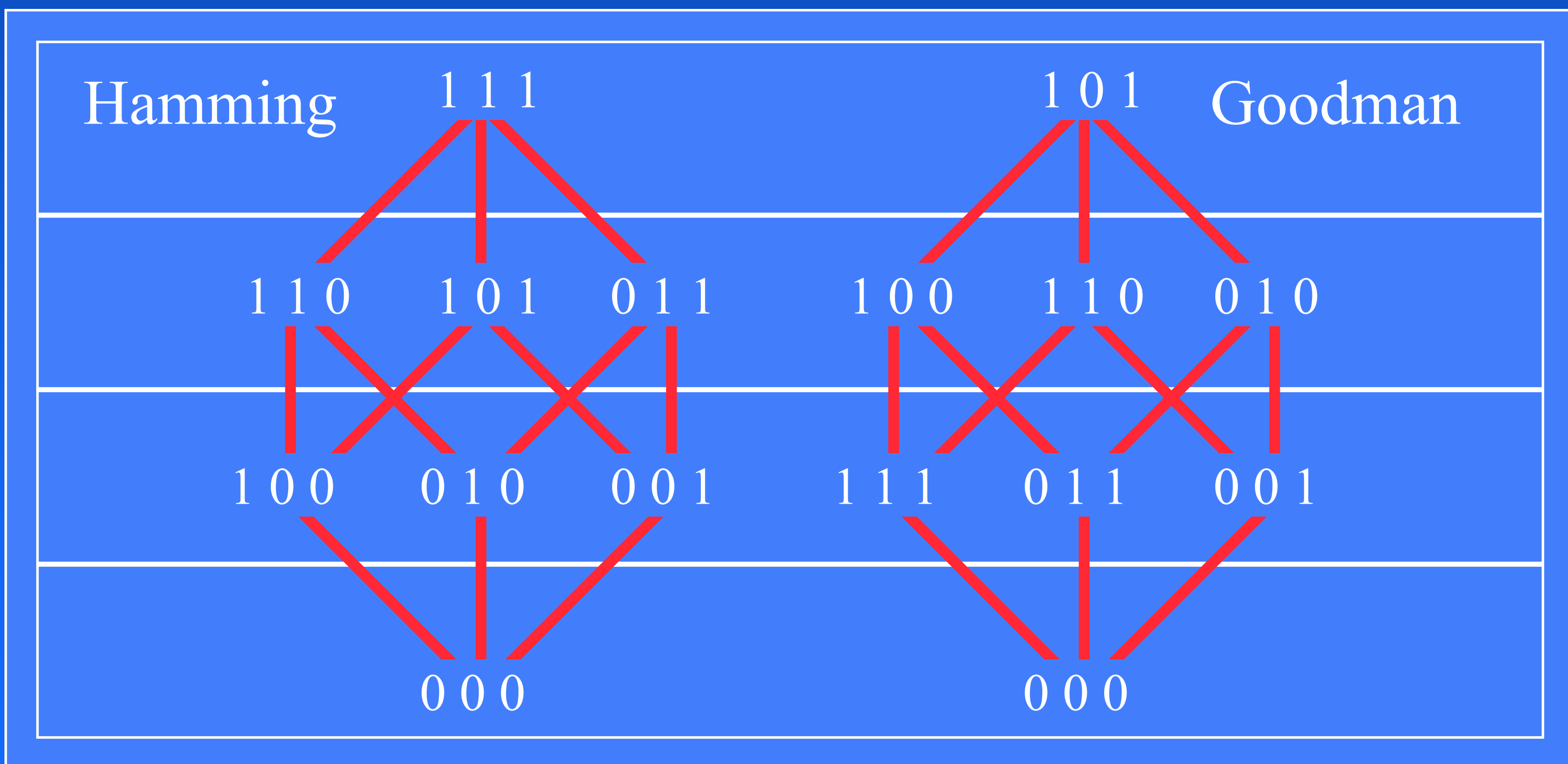
Wrench In the Works

- So b never rises above a when a is true (positive invariance)
- Now a and b agree forever, so can never be separated.
- So never converges in a or forgets refutation of b.

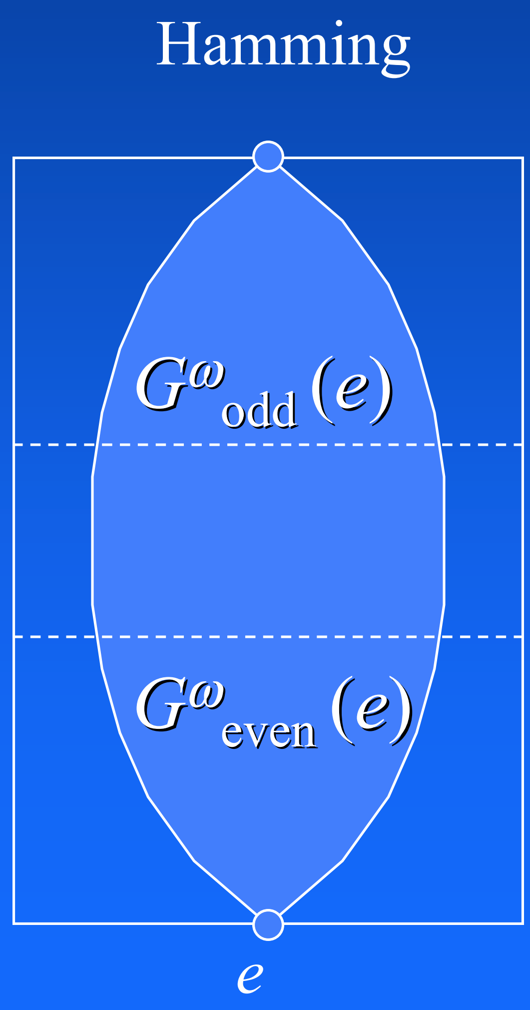


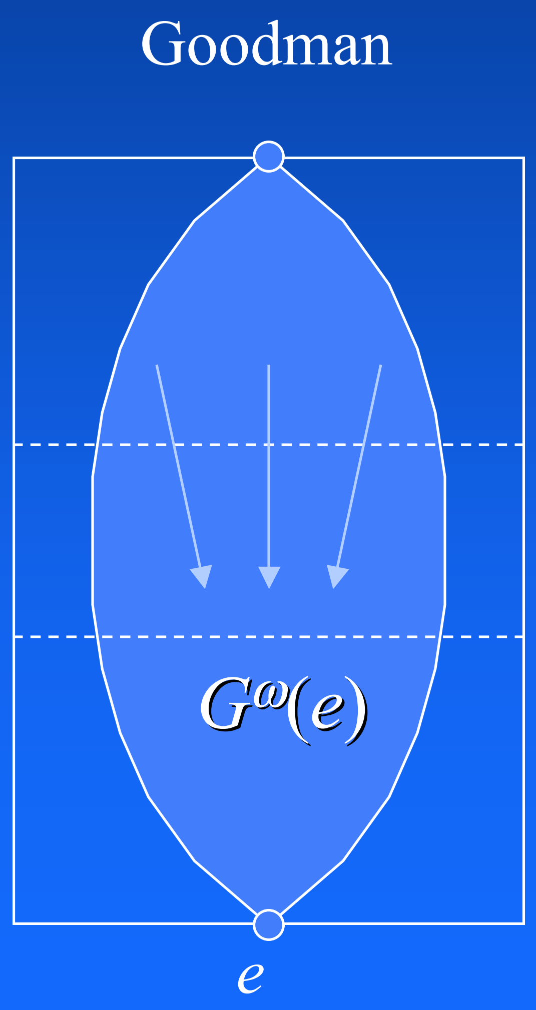
Hamming vs. Goodman Algebras

- $a \leq_H b \text{ mod } e \Leftrightarrow a$ differs from e only where b *does*.
- $a \leq_G b \text{ mod } e \Leftrightarrow a$ grues e only where b *does*.



Epistemic States as Boolean Ranks

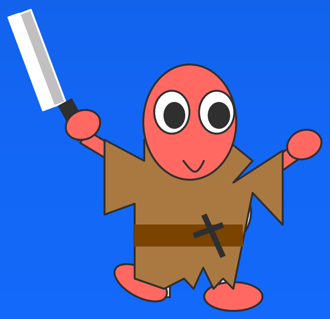




$*_{J,2}$ can identify $G^{\omega}(e)$

- *Proof*: Use the Goodman ranking as initial state
- Then $*_{J,2}$ always believes that the observed grues are the only ones that will ever occur.

- Note: Ockham with respect to reversal counting problem.



Classification: arbitrary grues

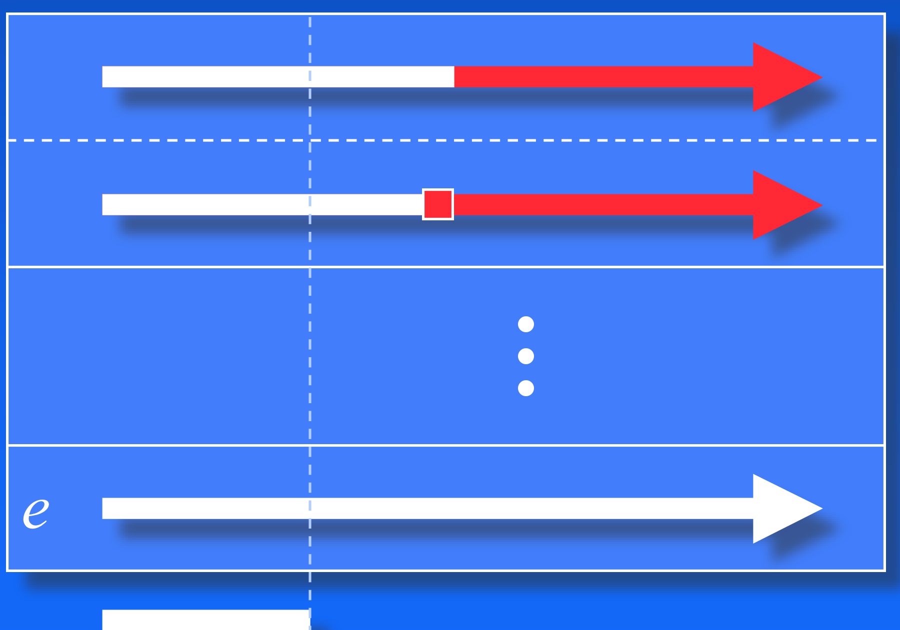
	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}(e)$	*no*	$\alpha = \omega$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{3}(e)$	*no*	$\alpha = n+1$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{2}(e)$	*no*	$\alpha = 3$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{1}(e)$	*no*	$\alpha = 2$	$\alpha = 2$	$\alpha = 1$	*yes*	*yes*
$G^{0}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

Methods $*_{\mathrm{J},1}$; $*_{\mathrm{M}}$ Fail on $G^1(e)$

- *Proof*: Suppose otherwise
- Feed e until e is uniquely at the bottom

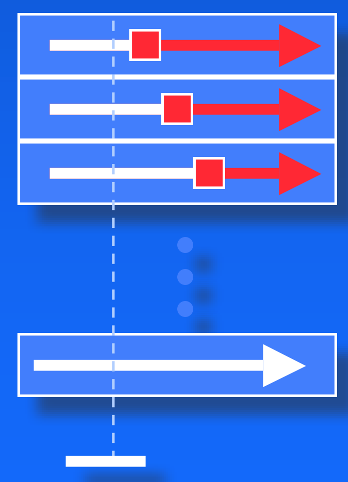
Methods $*_{J,1}$; $*_M$ Fail on $G^1(e)$

- By the well-ordering condition,



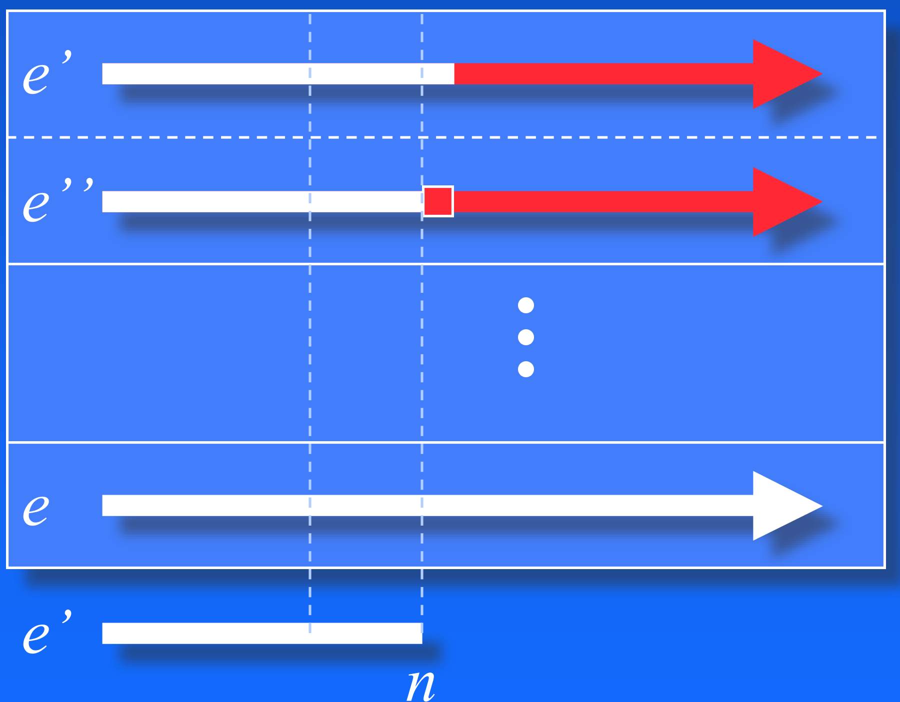
data so far

...else infinite descending chain



Methods $*_{J,1}$; $*_M$ Fail on $G^1(e)$

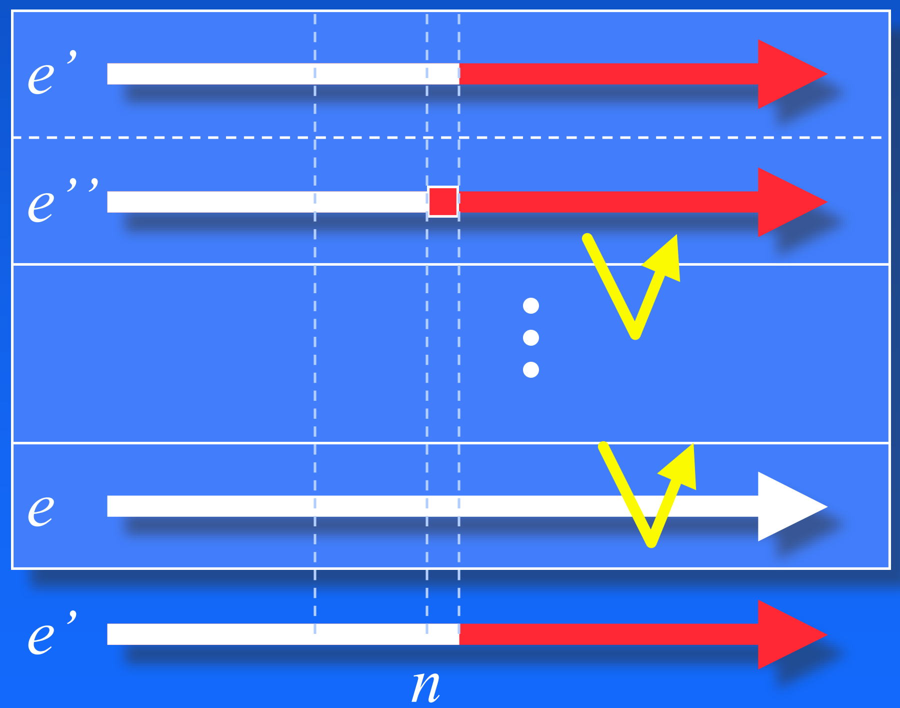
- Now feed *e'* forever
- By stage *n*, the picture is the same



positive order invariance

timidity and stubbornness

Methods $*_{J,1}$; $*_{M}$ Fail on $G^1(e)$



- At stage n +1, e stays at the bottom (timid and stubborn).
- So e' can't travel down (rule)
- e'' doesn't rise (rule)
- Now e'' makes it to the bottom at least as soon as e'

Classification: arbitrary grues

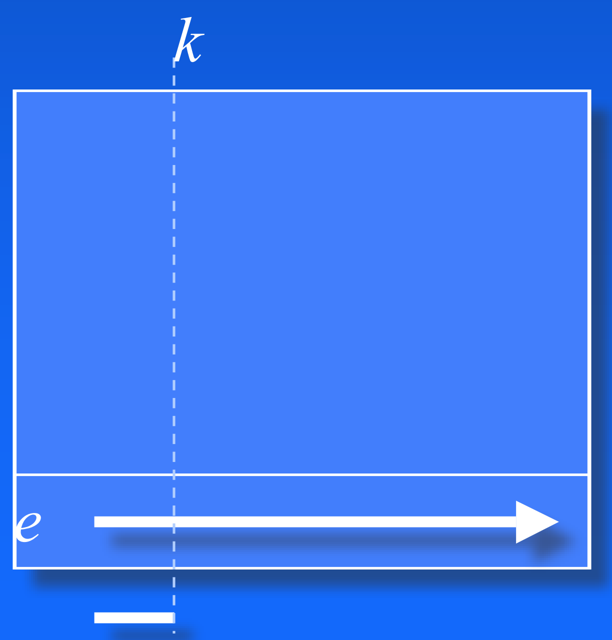
	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}(e)$	*no*	$\alpha = \omega$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
⋮				⋮		
$G^{3}(e)$	*no*	$\alpha = n + 1$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
⋮				⋮		
$G^{2}(e)$	*no*	$\alpha = 3$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{1}(e)$	*no*	$\alpha = 2$	$\alpha = 2$	$\alpha = 1$	*yes*	*yes*
$G^{0}(e)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	*yes*	*yes*

forced backsliding

Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

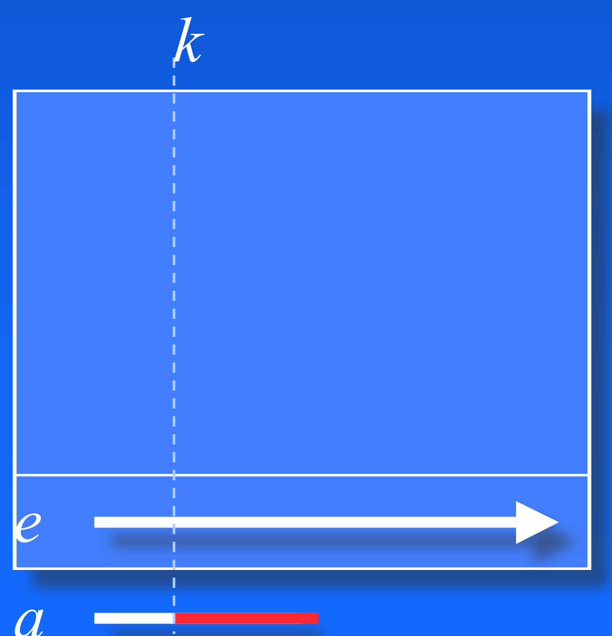
- *Proof*: Suppose otherwise
- Bring e uniquely to the bottom, say at stage k



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

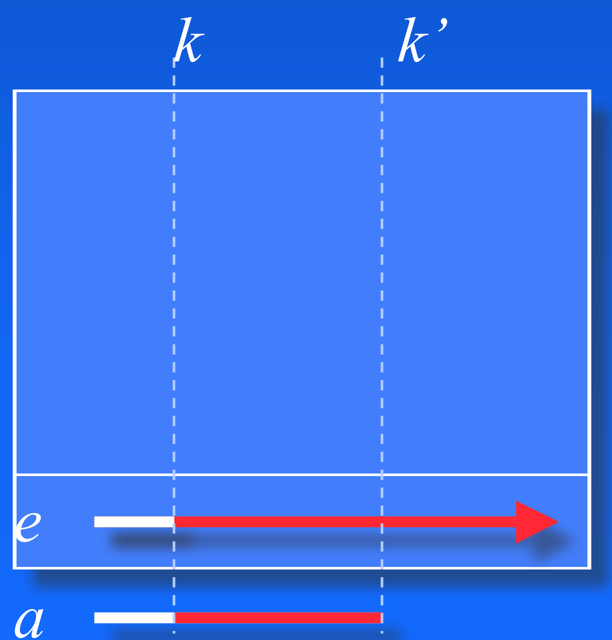
- Start feeding $a = e \ddagger k$



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

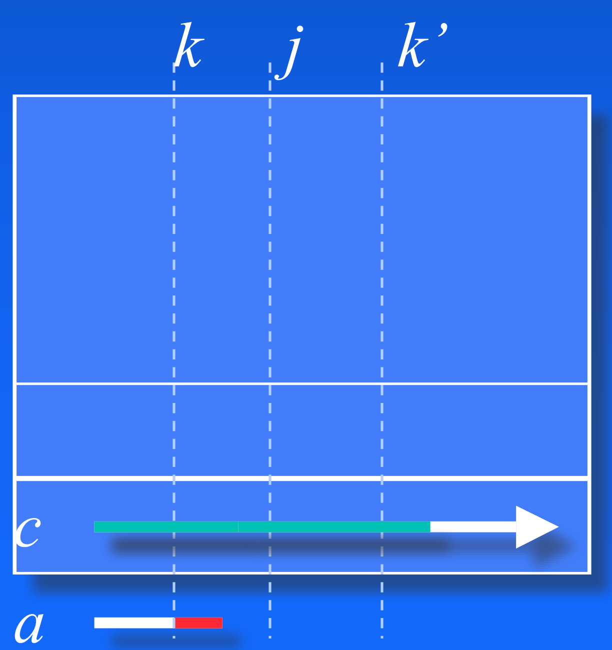
- By some stage k', a is uniquely down
- So between $k + 1$ and k', there is a first stage j when no finite variant of e is at the bottom



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

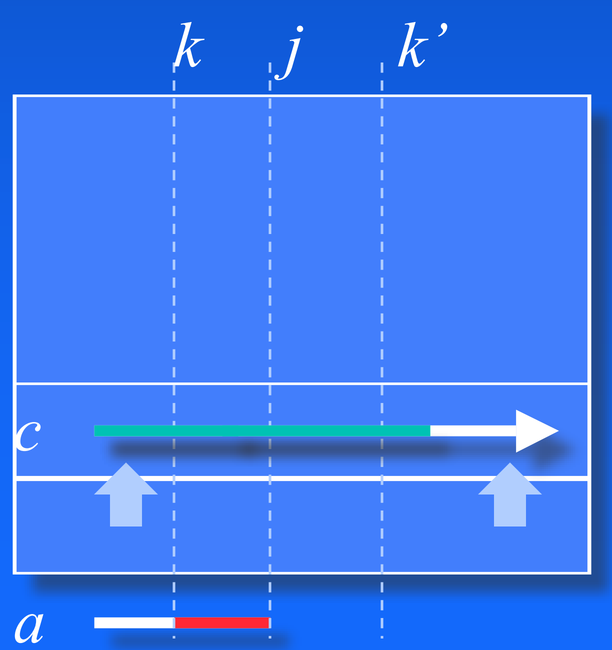
- Let c in $G^2(e)$ be a finite variant of e that rises to level 1 at j



Method $*_{R,1}$ Fails on $G^2(e)$

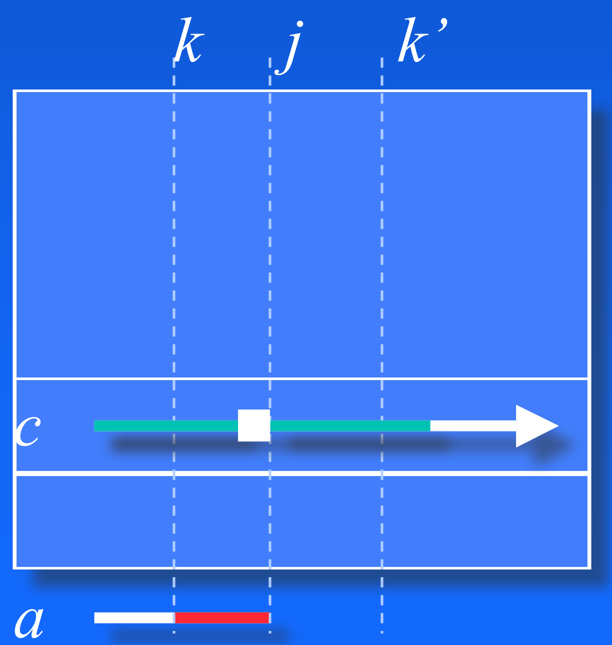
with Oliver Schulte

- Let c in $G^2(e)$ be a finite variant of e that rises to level 1 at j



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

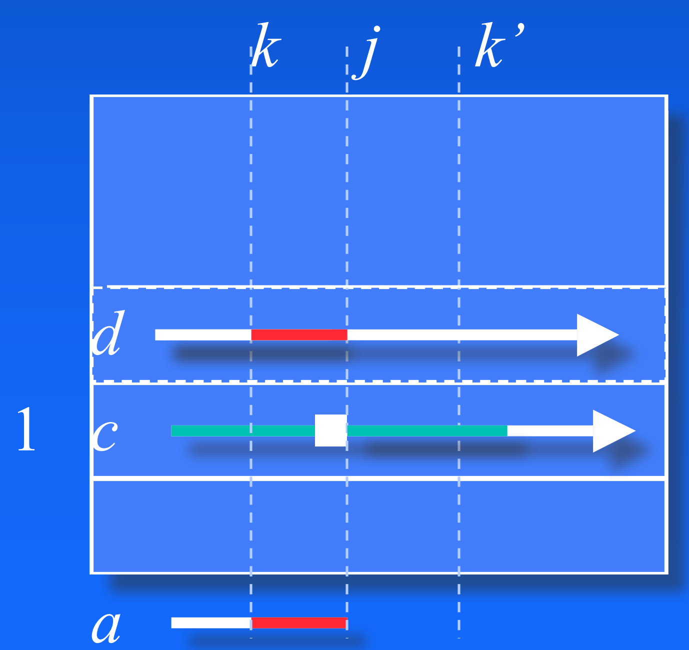


- So $c(j - 1)$ is not $a(j - 1)$

Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

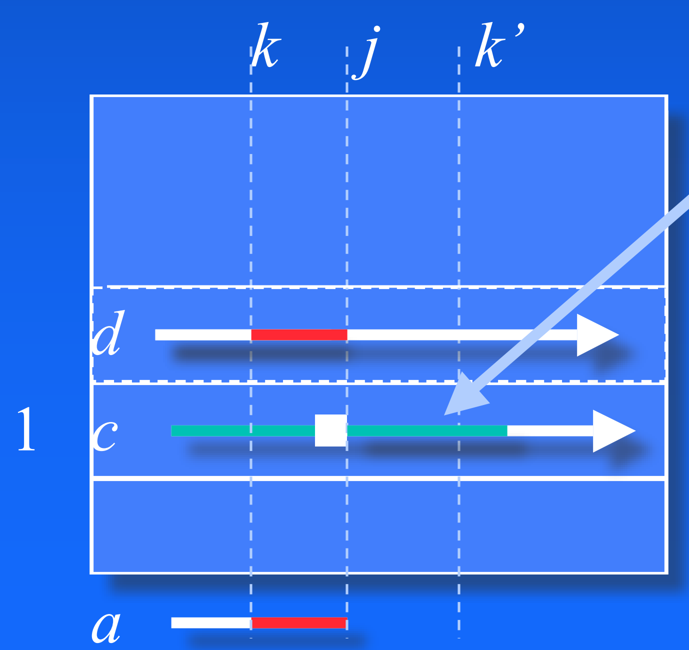
- Let d be a up to j and e thereafter
- So is in $G^2(e)$
- Since d differs from e, d is at least as high as level 1 at j



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

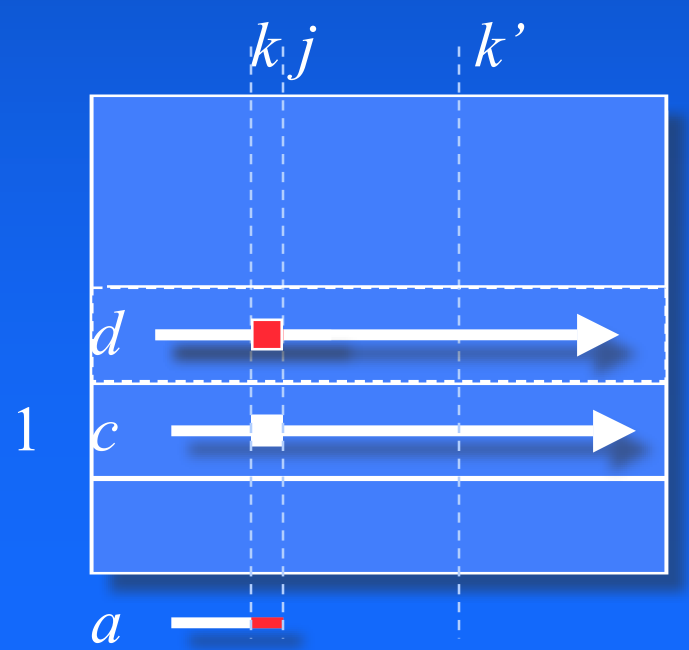
- Show: c agrees with e after j.



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

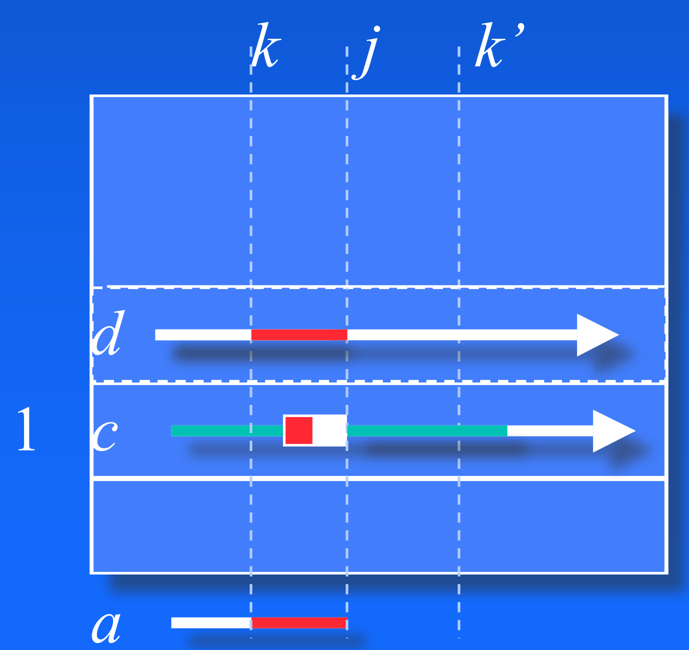
- Case: $j = k+1$
- Then c could have been chosen as e since e is uniquely at the bottom at k



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte

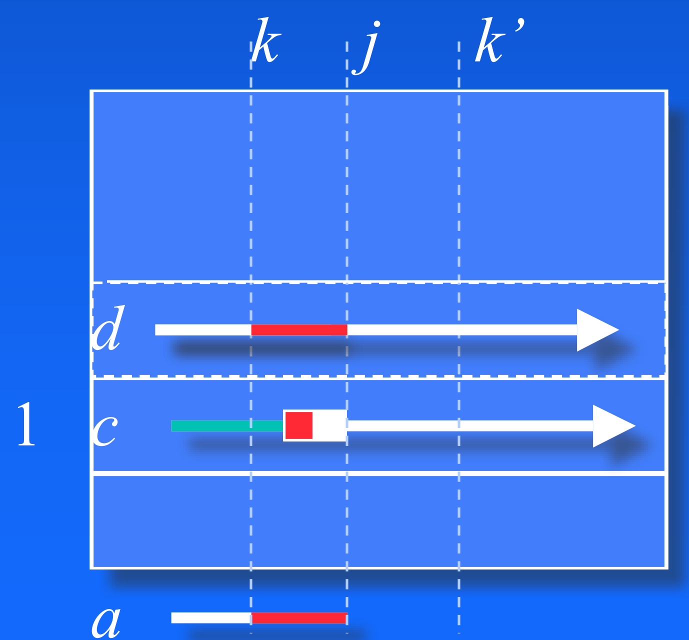
- Case: $j > k+1$
- Then c wouldn't have been at the bottom if it hadn't agreed with a (disagreed with e)



Method *R,1 Fails on $G^2(e)$

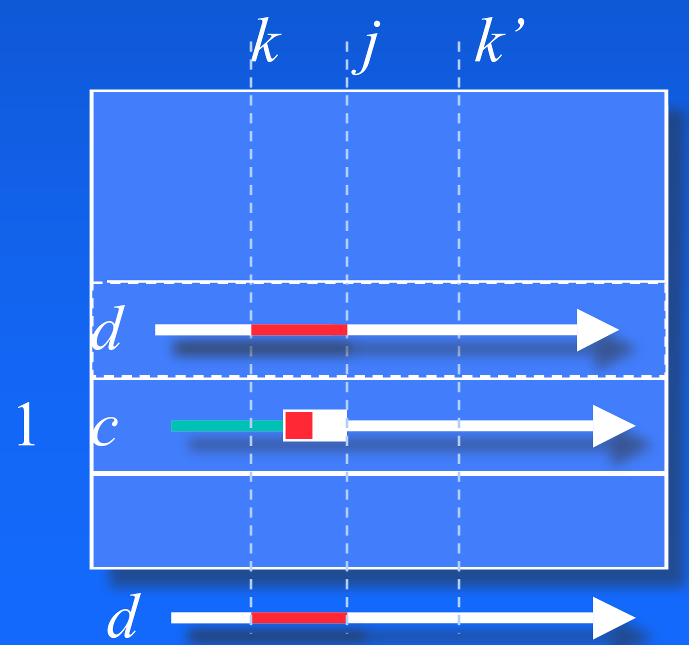
with Oliver Schulte

- Case: $j > k+1$
- So c has already used up its two grues against e



Method $*_{R,1}$ Fails on $G^2(e)$

with Oliver Schulte



- Feed c forever after
- By positive invariance, either *never projects* or *forgets* the refutation of c at j-1

Without Well-Ordering

	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{\omega}(e)$	*no*				*yes*	*yes*
				$\vdots$		
$G^{3}(e)$	*no*				*yes*	*yes*

infinite descending
chains can help!

	Min	Flush	Jeffrey	Ratch	Lex	Cond
$G^{2}(e)$	*no*				*yes*	*yes*
$G^{1}(e)$	*yes*				*yes*	*yes*
$G^{0}(e)$	*yes*				*yes*	*yes*

Summary

- Belief revision constrains possible inductive strategies
- "No induction without contradiction" (?!!)
- "Rationality" weakens learning power of ideal agents.
- Prediction *vs*. memory
- Precise recommendations for rationalists:
  - boosting by 2 *vs*. 1
  - backslide *vs*. ratchet
  - well-ordering
  - Hamming vs. Goodman rank