Backward-then-forward planning in the cerebrum

From maze runs to hierarchical map, and back again

Edward Hedgecock
Department of Biology
Johns Hopkins University

11 May 2022

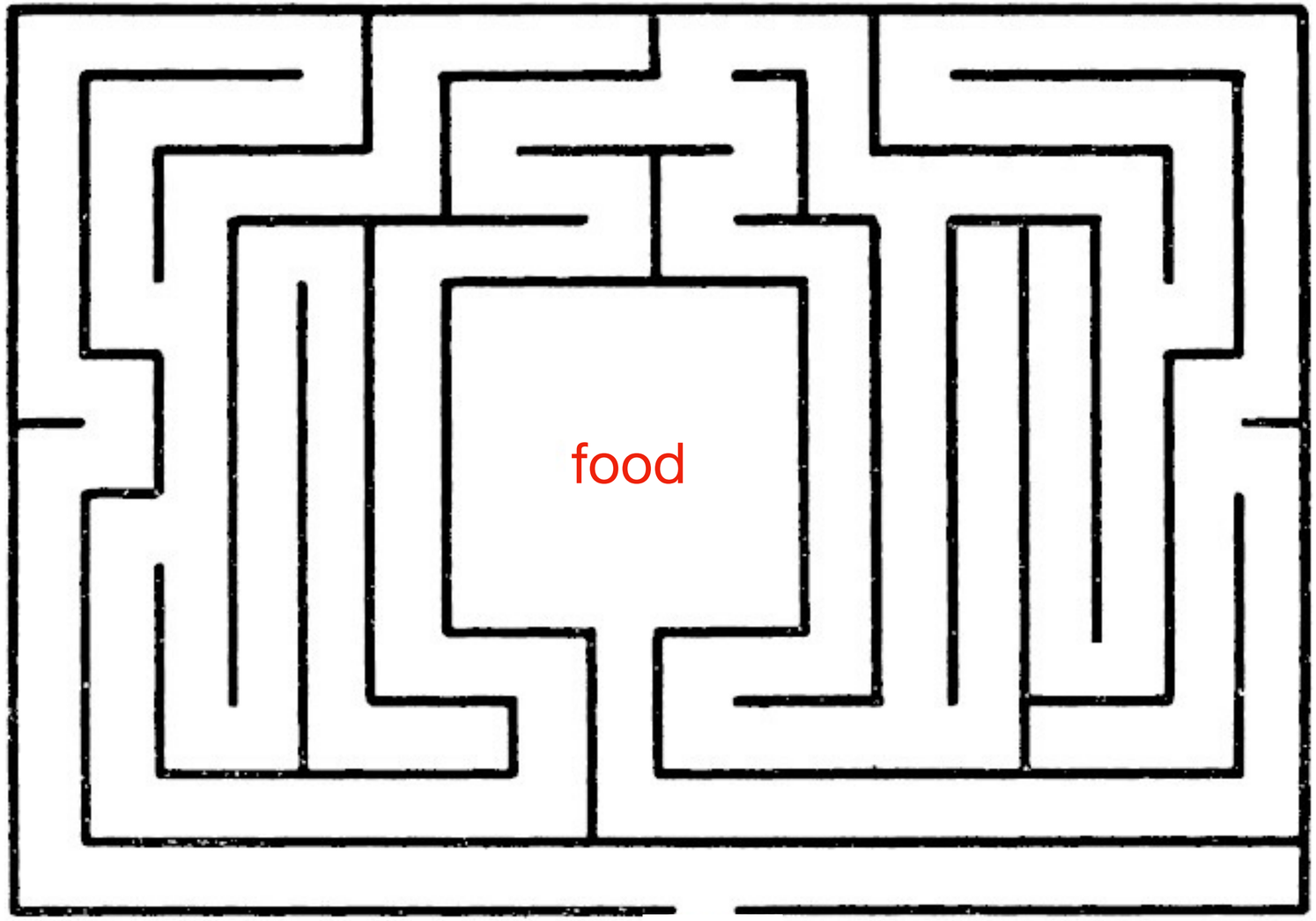
Is there a fundamental problem of motivated behavior?

Duality of explore & exploit

EXPLORE	prospect novel deposits	curiosity novel causes	discover novel places	map novel maze
EXPLOIT	mine familiar deposits	habit familiar effects	revisit familiar places	navigate map & maze

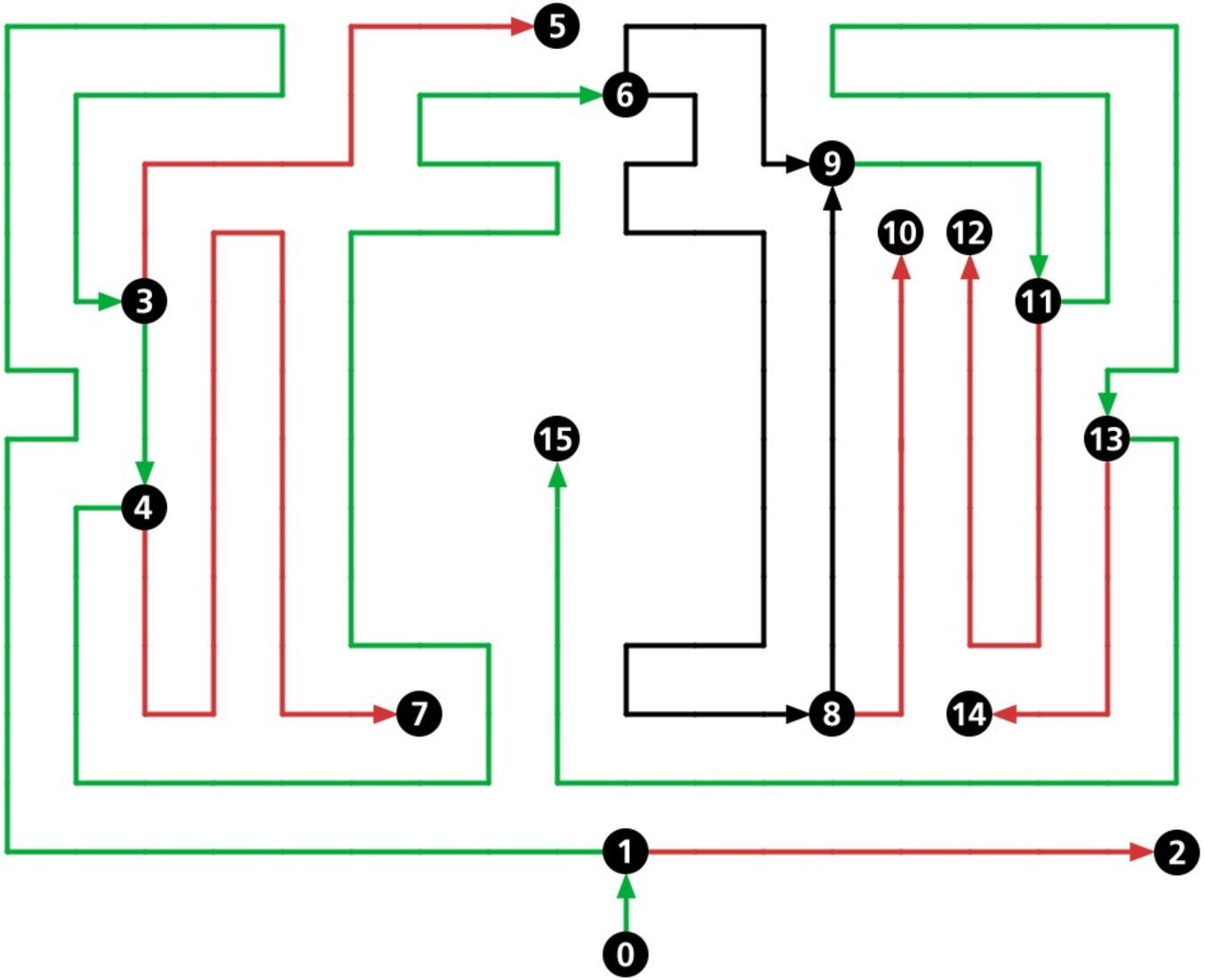
Experimental study of the mental processes of the rat

Willard Small (1901). *American Journal of Psychology*



Experimental study of the mental processes of the rat

Willard Small (1901). *American Journal of Psychology*



Maze behavior of laboratory rats

Motivation

- Present need (*viz* hunger)

Initial exploration

- Trial-and-error search

Tentative solutions with repeated trials

- Speed increases on runways
- Hesitation & errors decrease at choice points

Habitual solution with 'over-training'

- Automatic movements along fixed path to food
- Favor shorter path, over longer one
- 'Error-free' runs of true path of 40 feet in under 20 seconds

Natural behavior of wild rats

Motivation

- Cautious curiosity about resources & dangers
- Investment in future needs, not just present ones

Continual exploration interleaved with exploitation

- Explore novel paths for undiscovered shortcuts & resources
- Monitor familiar paths for emergent shortcuts, obstacles, resources & dangers

Home-base behaviors (grooming, rearing, circling)

- Starting place, or another place with natural protection

Outward exploration (foraging)

- Slow steps & circuitous paths, avoiding open spaces & following walls
- Revisit familiar places & discover new ones

Approach & avoid behaviors (appetite, aversion)

- Sample food for taste & wholesomeness
- Defense (freeze, flee or fight) & post-threat (vigilance & place aversion) behaviors

Homeward run by a rapid beeline path

The determiners of behavior at a choice point

Edward Tolman (1938). *Psychological Review*

Latent learning

- Law-of-effect explained trial-and-error learning from rewarded runs
- Not latent learning from exploring the maze before food is added
- Nor selecting alternative path to water when thirsty

Route planning (insight)

- Detour around new obstacle
- Shortcut through new passage

Tolman's take-home message

- Learning & reasoning seen best by changing means-ends problem

Cognitive maps in rats & men

Edward Tolman (1948). *Psychological Review*

Initial exploration, narrow feature-poor map

- Single path from start to finish
- Attend few features

Routine exploitation, broad feature-rich map

- Multiple paths between any two places on the map
- Attend many features

Paradox of habit

- Cognitive map broadens from new experience
- Habitual paths narrow with repeated practice

Habit formation

Two kinds of memory, or one kind with several levels

Hand-off from flexible plans to rigid habits

- From motivated behavior to automatic behavior (James 1890)
- From place learning to response learning (Tolman & students 1946-48)
- From locale system to taxon system (O'Keefe & Nadel 1978)

Downward distribution of plans within hierarchical map

- Plans narrow even as map broadens
- Use-dependent memory consolidation to lower levels
- Use-dependent program compilation to lower levels

My research & today's talk

Icebergs, intuitions & a few results

Formulate the problem & what a solution might look like

- Boolean mazes
- Hierarchical maps
- Plans with subplans

Learning for reasoning

- Minimally curious student (*cf* active learning Angluin 1987)

Reasoning for learning

- Maximally forgetful professor (*cf* bisimulation Park 1981)
- Backward-then-forward planning & forward-then-backward parsing

What of the cerebrum?

- Formal explanation of psychological & neural phenomena
- Neural realization of formal principles, *viz* maps, plans, rewrite rules

Is there a fundamental problem of motivated behavior?

Forward & backward search

MAPPING	forward in maze	novel goal	discover path
PLANNING	backward in map	familiar goal	plan path
ACTING	forward in map & maze	test current state	move next state

Means-ends reasoning

From Aristotle to STRIPS planning

PLANNING

backward search
in map

Initial state
last marked
& most tests

Goal state
first marked
& fewest tests

ACTING

forward path
in map & maze

Initial state
first reached

Goal state
last reached

Boolean mazes ✓

Taxonomy of finite processes

States	$s \in S$	✓	states	$\{s_0, s_1, s_2 ...\}$	all processes
			initial	s_0	games automata
Events	$P \subseteq S \times S$	✓	moves	{P, Q, R ...}	all processes
	$U \subseteq S \times S$	✓	co-moves	{U, V, W ...}	games
Tests			none		transition systems
	$S \to \{0,1\}$		Boolean	win / lose accept / reject	games Kleene automata
	$S \to \{0,...k\}$		multivalued		Moore automata
$A \subseteq S$	$A \to \{0,1\}$	✓	Booleans	{A, B, C...}	Kozen automata

Boolean maze

Moves & tests

Empty moves

0 R 0

Commuting moves

0 P 1 Q 4

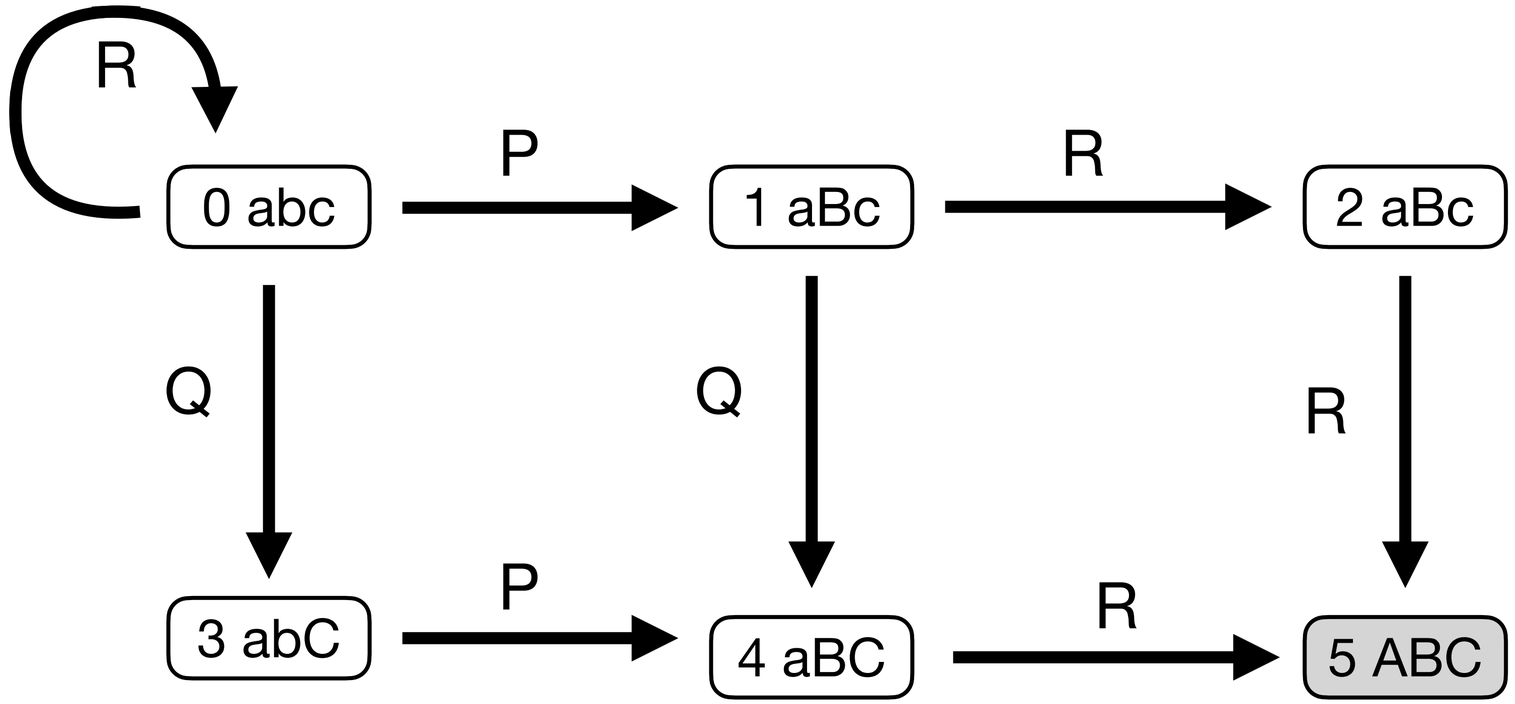
0 Q 3 P 4

Alternative paths

0 P 1 Q 4 R 5

0 P 1 R 2 R 5

0 Q 3 P 4 R 5



Distance in maze? In map?

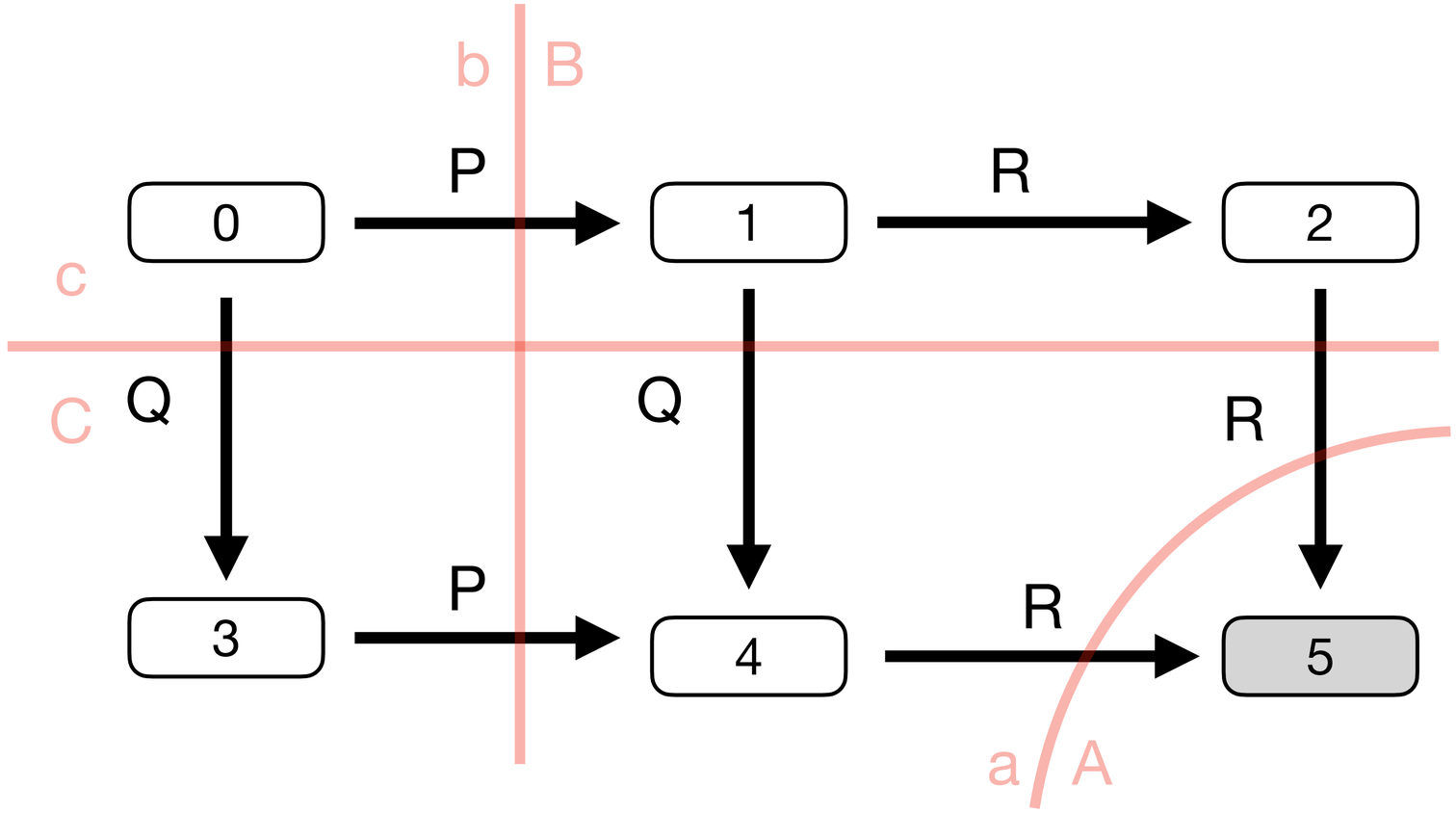
Metrics of moves or tests

Shortest path

- move-based *viz* egocentric
- length of shortest known path

Boolean difference

- test-based *viz* allocentric
- sum of known differences



Moves & co-moves

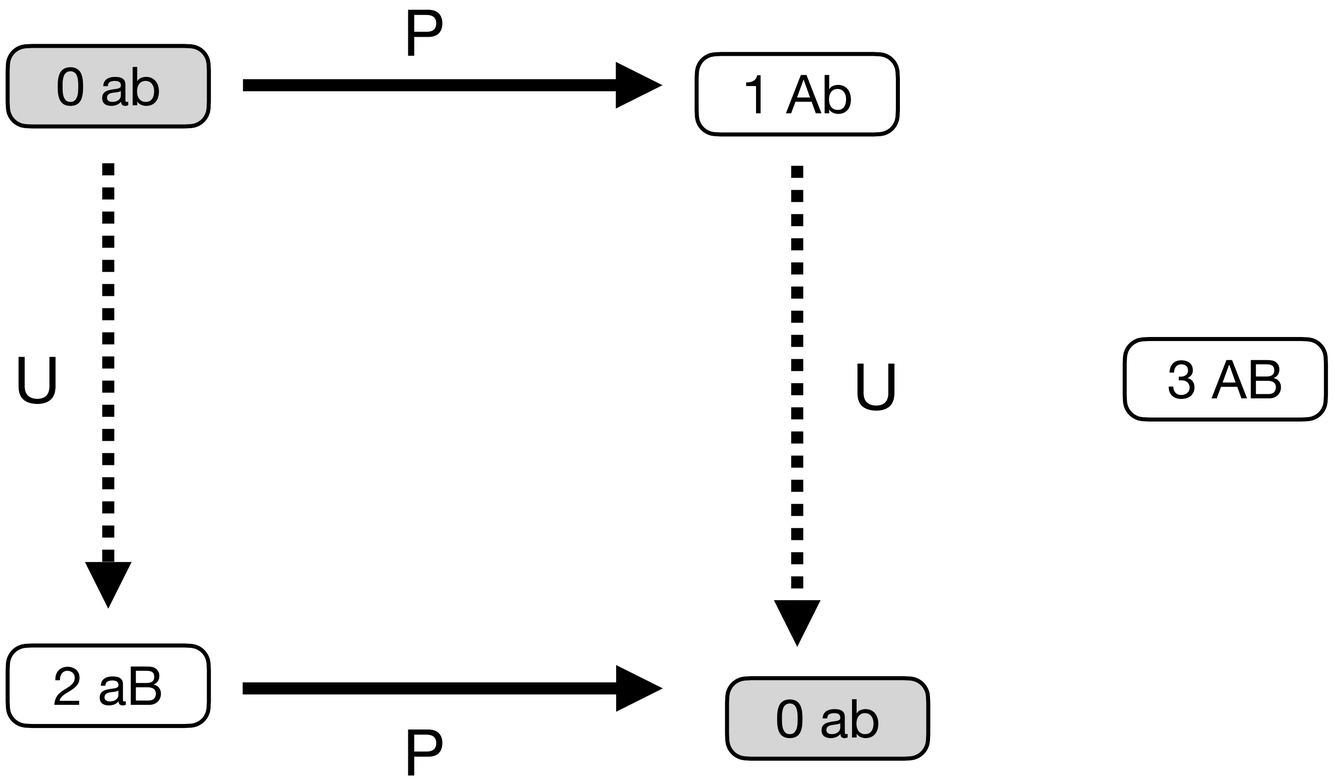
example : producer-consumer maze, or homeostat

States

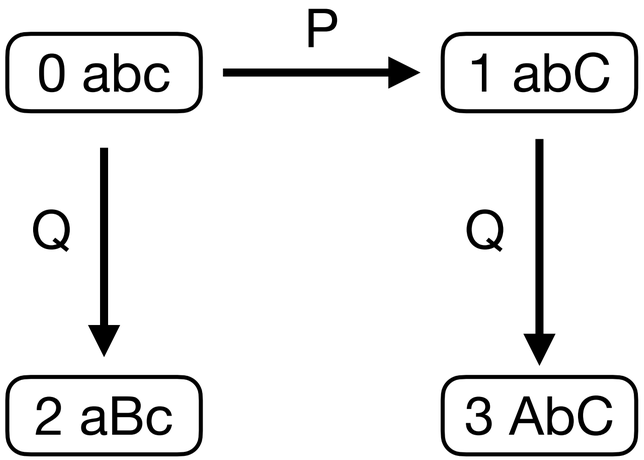
- ab just right
- Ab too much
- aB too little
- AB unreachable

Paths

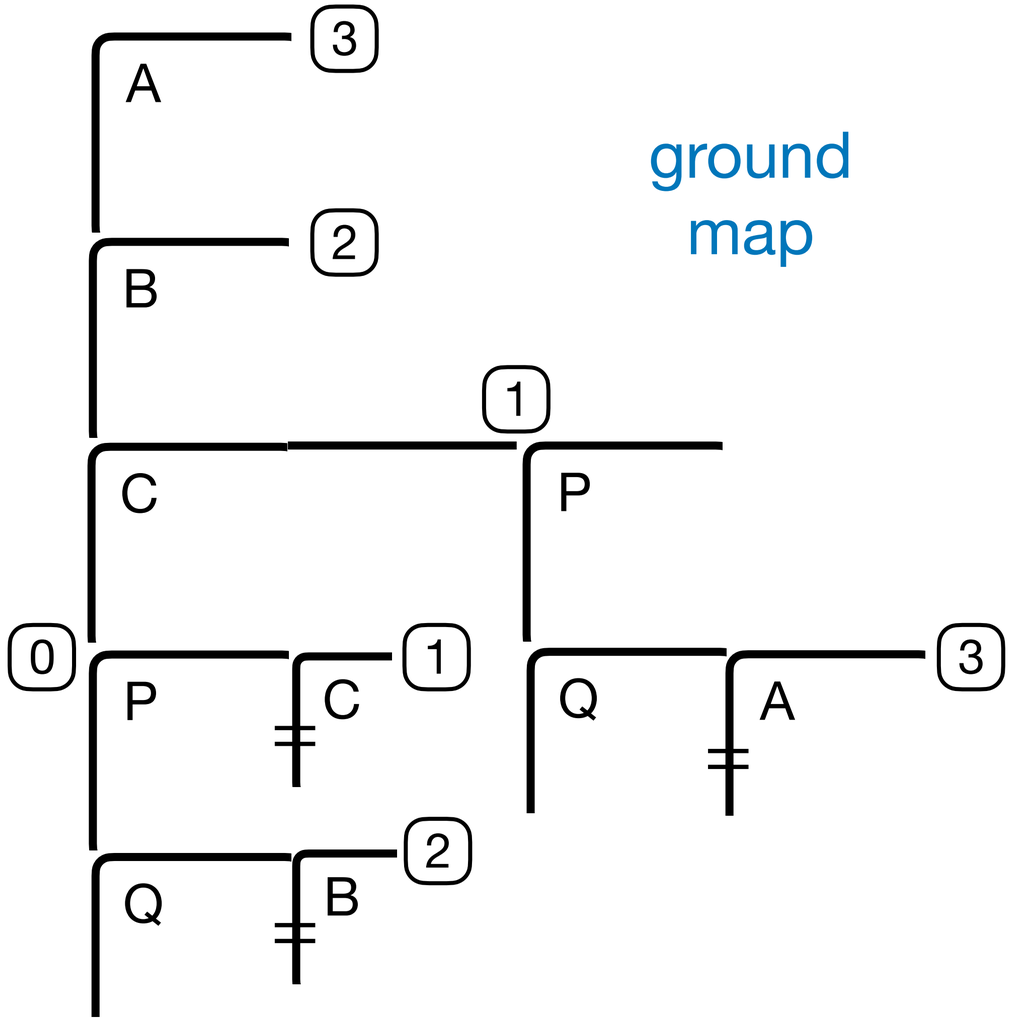
- 0 P 1 U 0
- 0 U 2 P 0

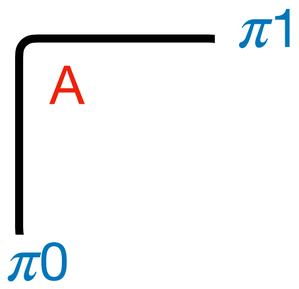


Boolean maze

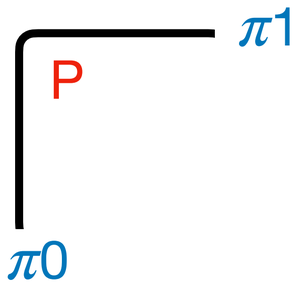


ground map

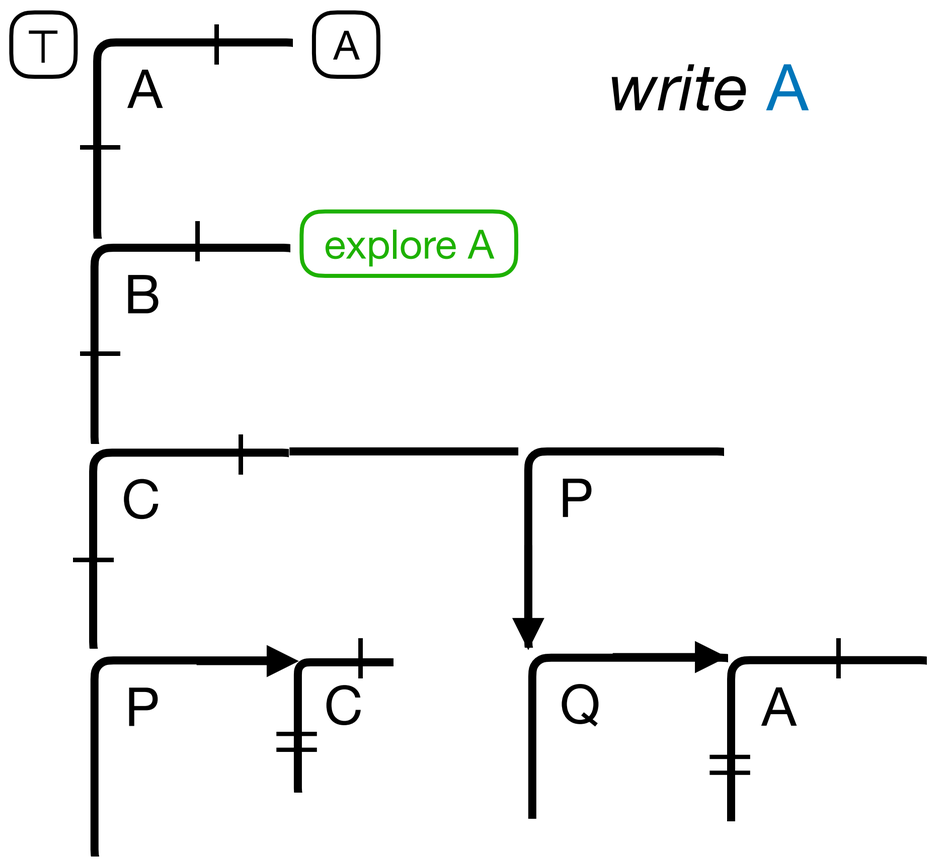





```
test A; π1 + test a; π0
  if A then π1 else π0
```

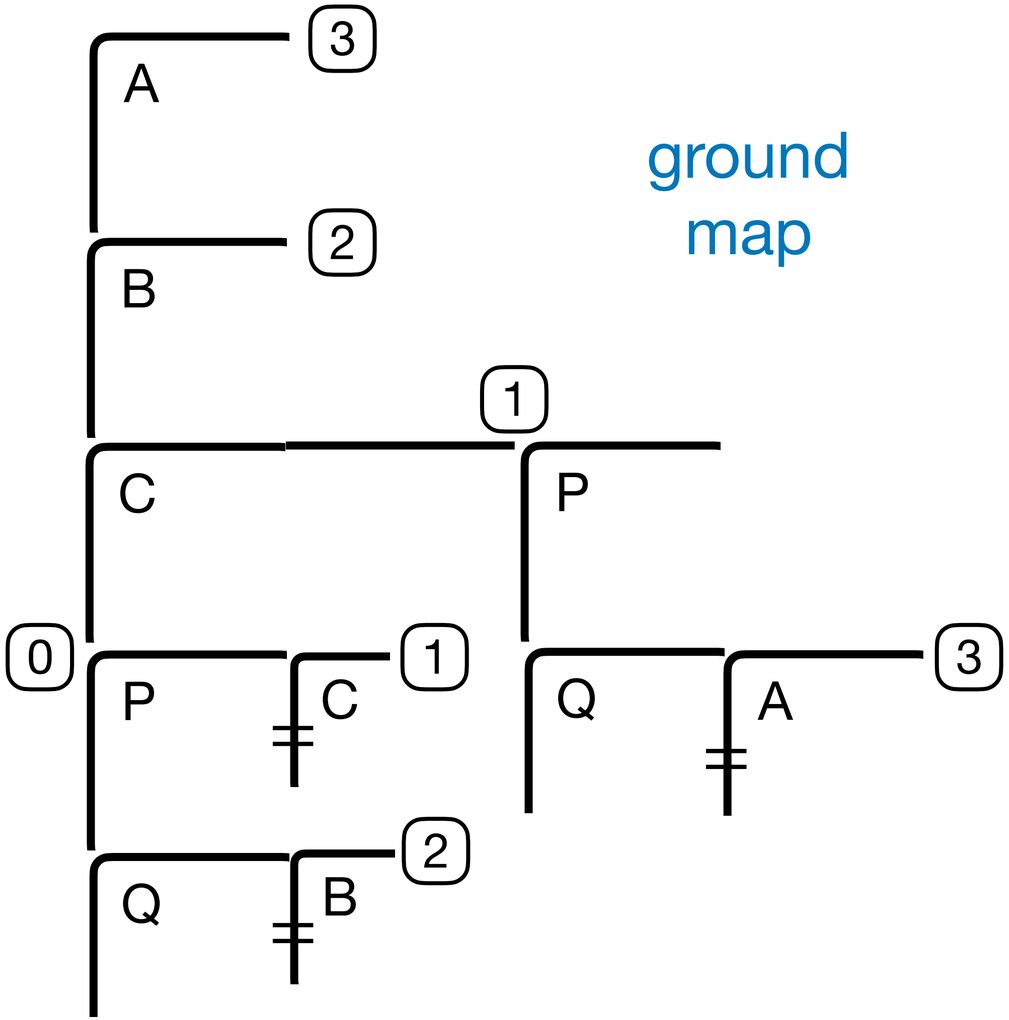


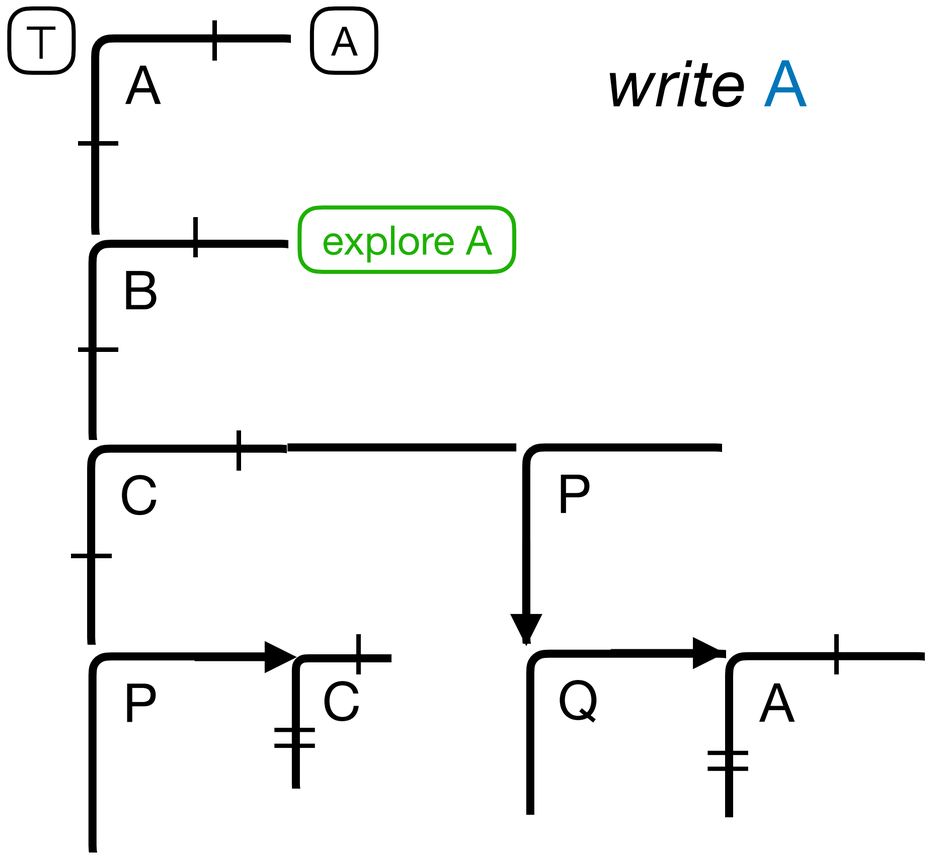

```
move P; π1 + move p; π0
move P; π1 + skip P; π0
```



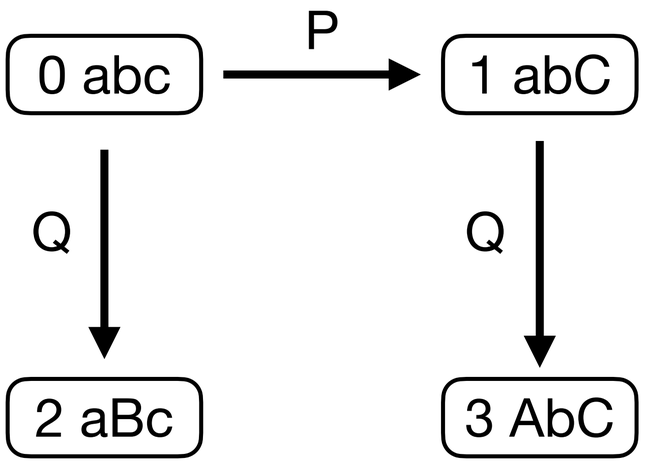
write A

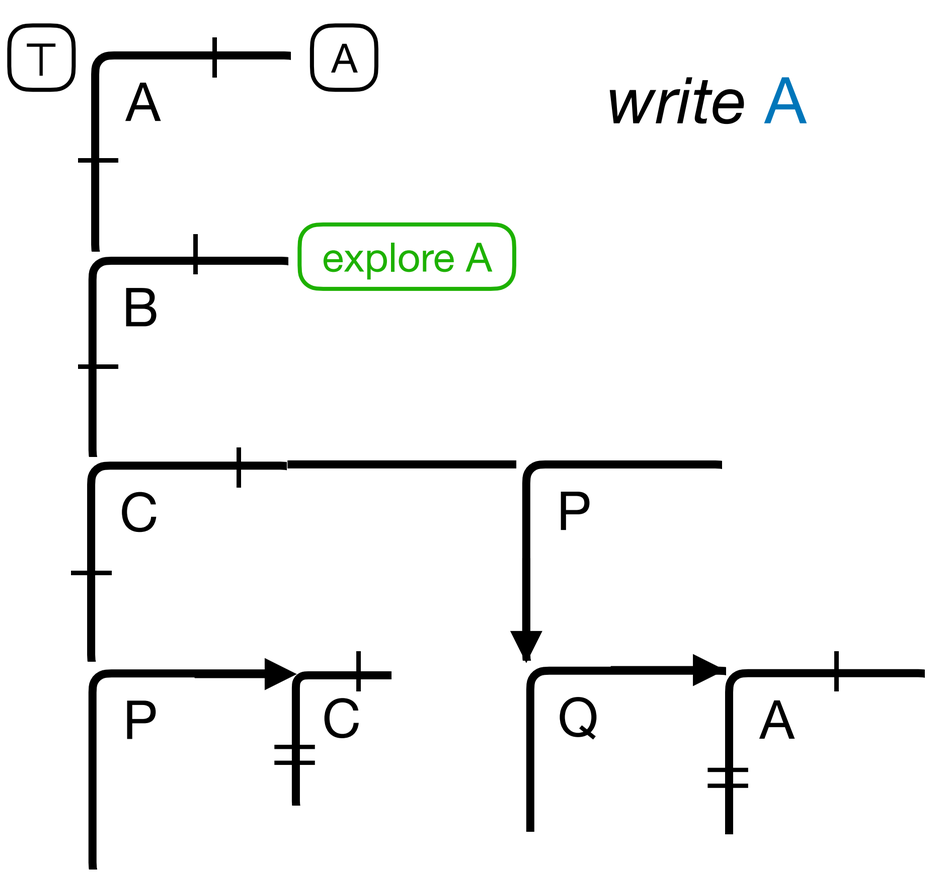
ground map



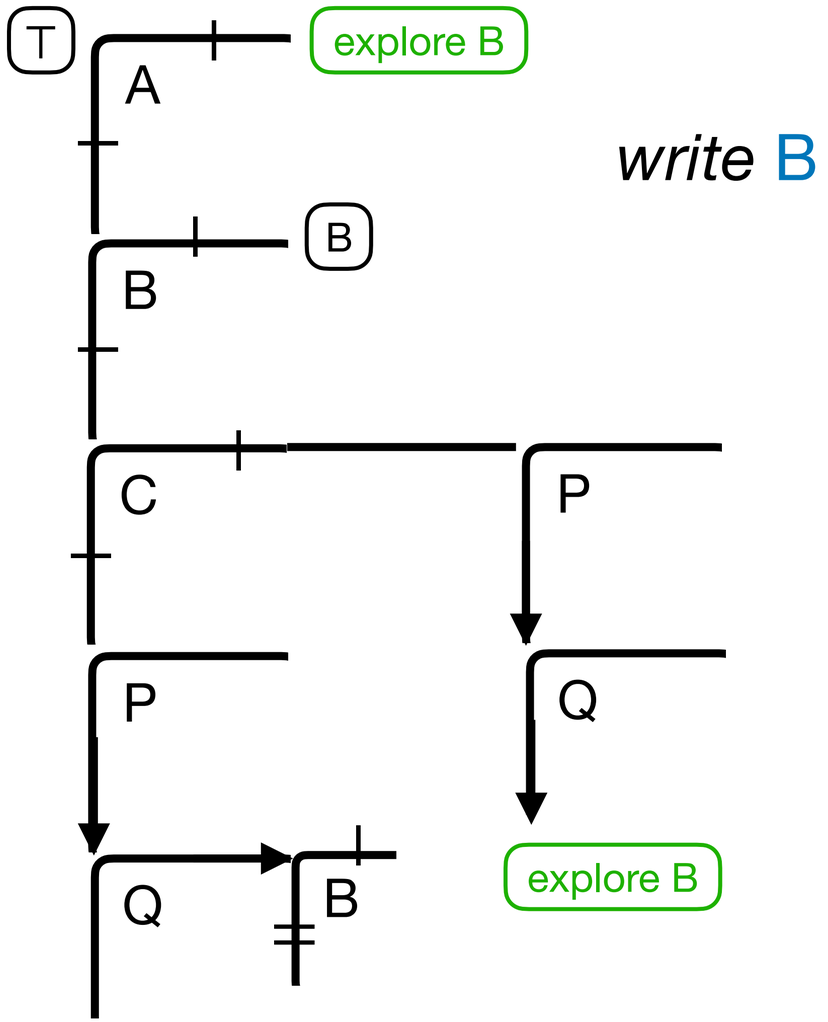


Boolean maze

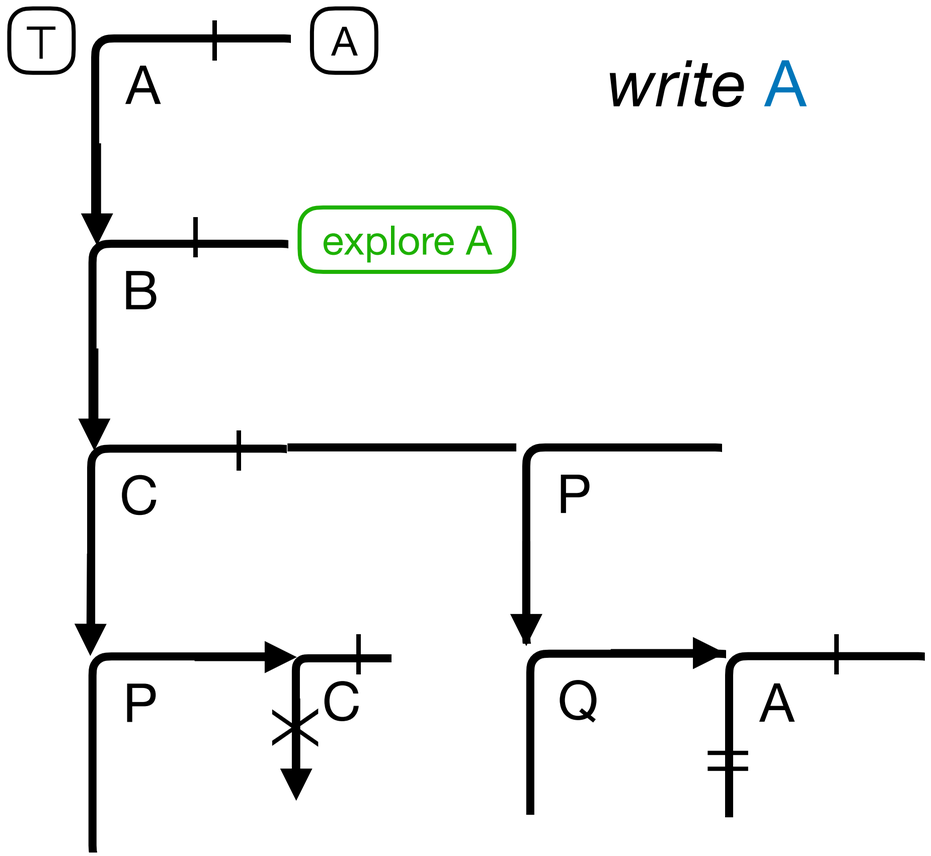




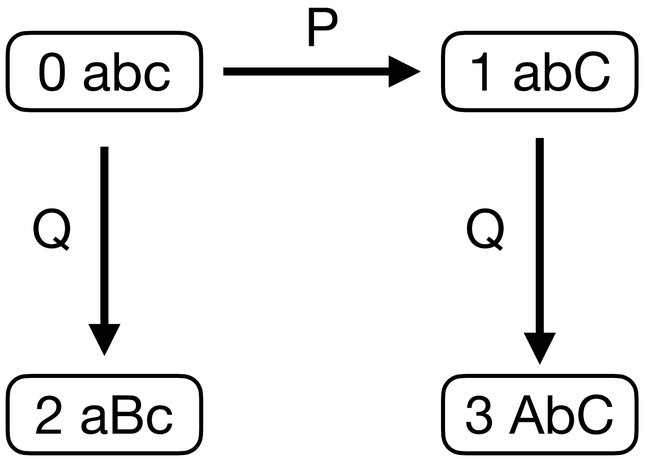
write A



write B



Boolean maze



When a plan screws up!

The contract is breached at a CWA

Plans and the Structure of Behavior

Miller, Galanter & Pribram (1960)

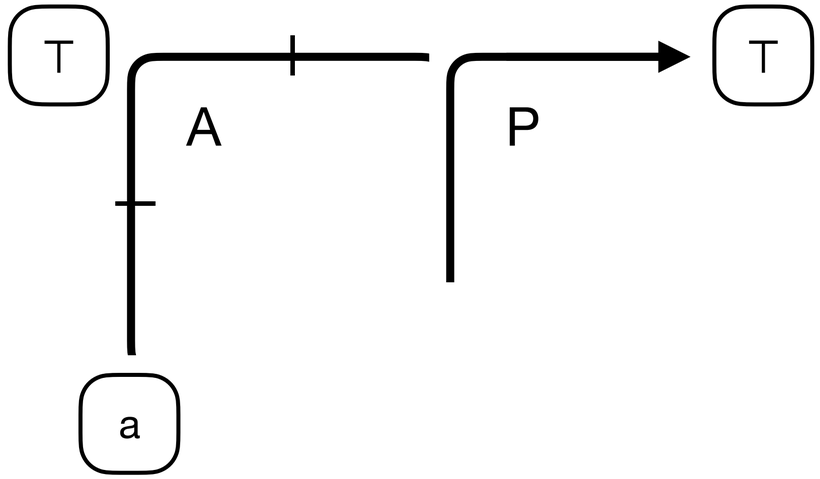
TOTE units

test-operate-test-exit

While programs

while A *do* P

TOTE unit



An axiomatic basis for computer programming

C.A.R. Hoare (1969). *CACM*

Antoine de Saint-Exupéry

"A goal without a plan is just a wish"

Tony Hoare

"A plan without a goal is just a program"

Partial correctness assertion, or contract

{A} P {B}

{**assume**} **program** {**assert**}

Relation algebra (Tarski & Givant 1987)

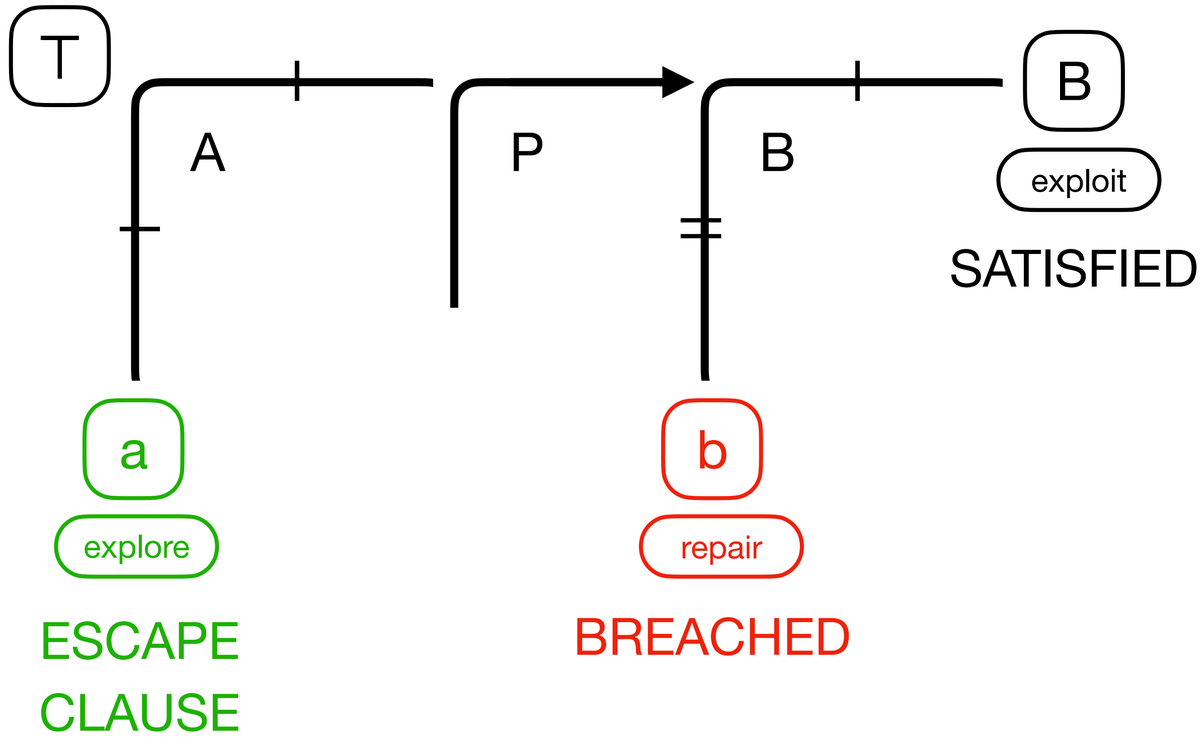
$(A \times b) \cdot P = 0$

breach · **program** = 0

Kleene algebra with tests (Kozen 2000)

APb = 0

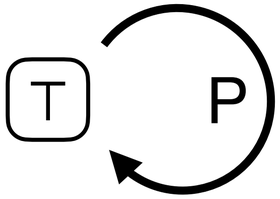
$\overline{\Diamond}$ APb

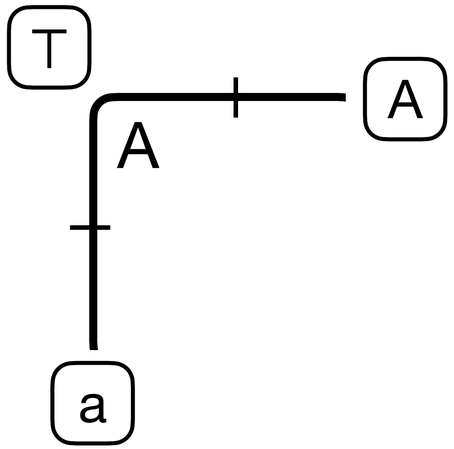


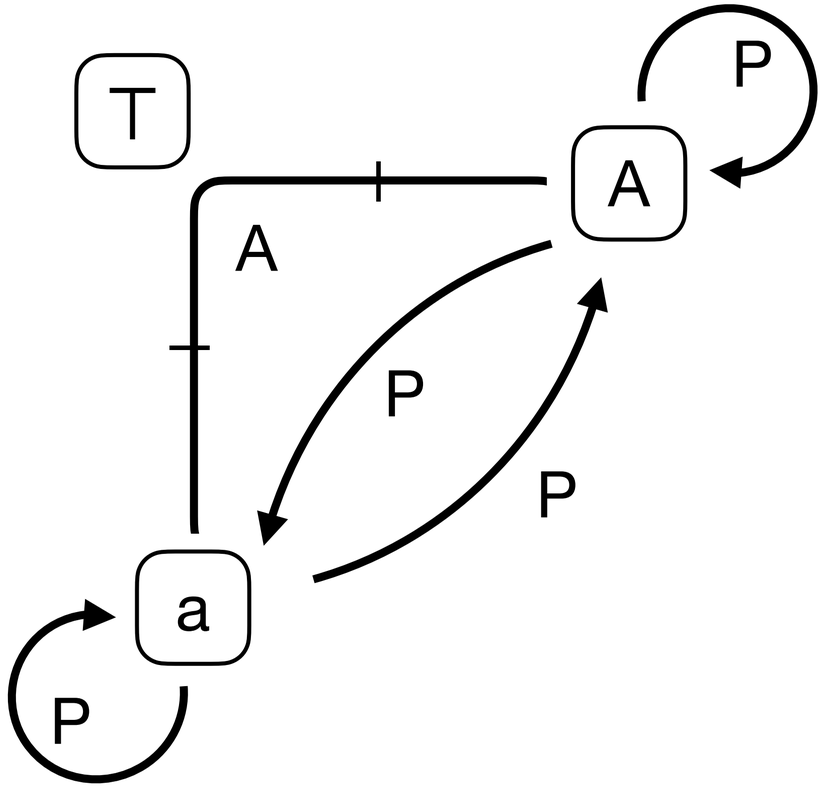
Every move & test specifies four plans

stay off + *turn* off + *stay* on + *turn* on

write off + *write* on





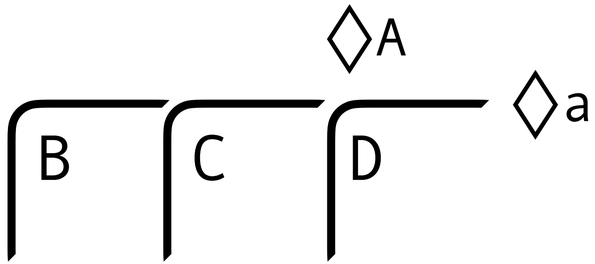


{T}P{T} = {a}P{a} + {A}P{a} + {A}P{A} + {a}P{A}

= {T}P{a} + {T}P{A}

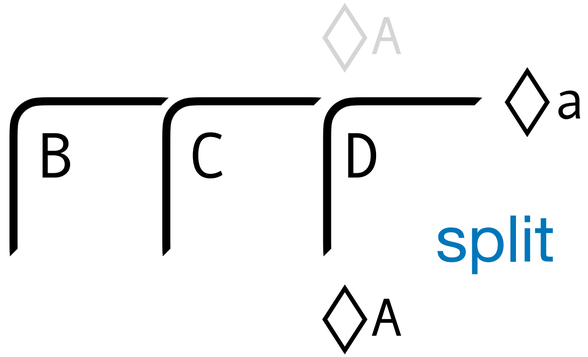
Decision tree learning

Minimally curious student



◊ABC

◊aBCD

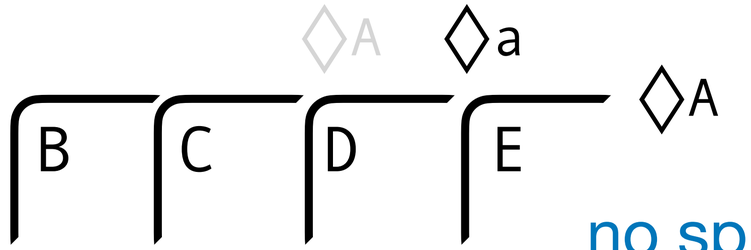


split

◊ABC

◊aBCD

◊ABCd



no split
'leap frog'

◊ABC

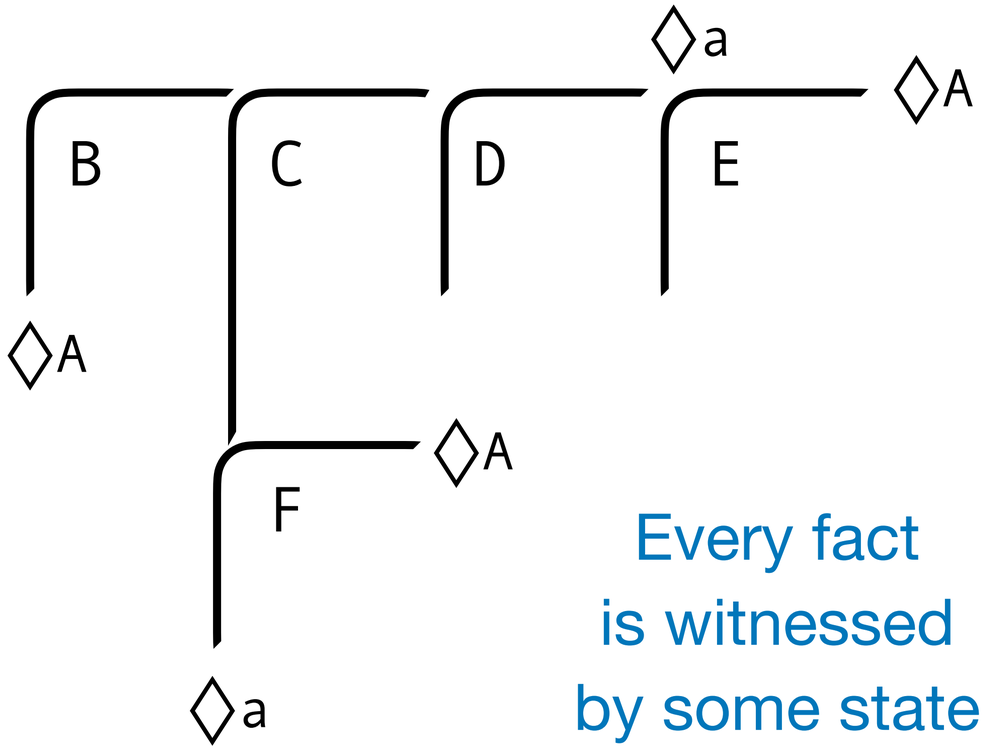
◊aBCD

◊ABCDE

Closed world assumption

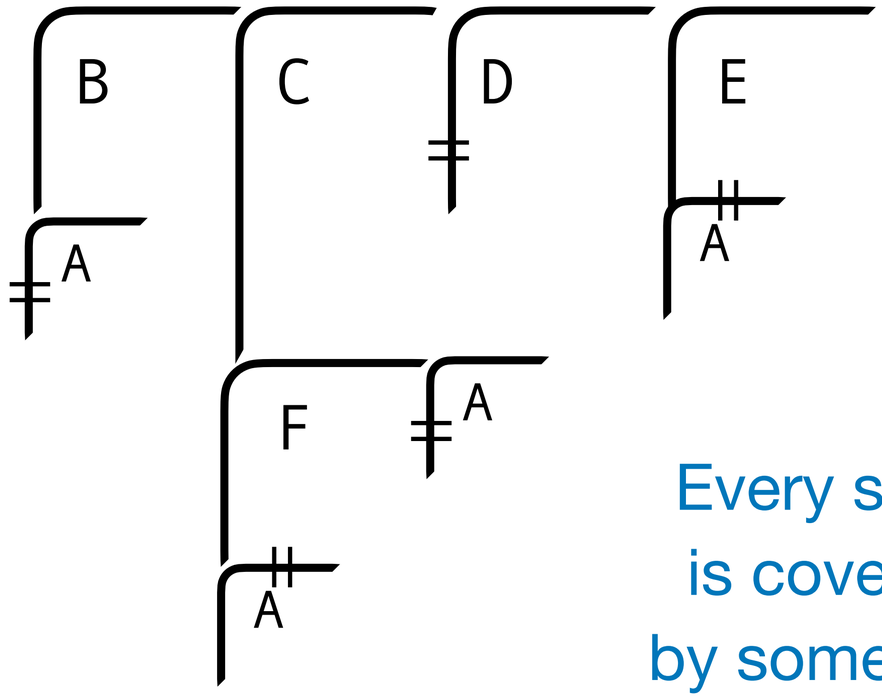
"absence of evidence is evidence of absence"

Facts



Every fact
is witnessed
by some state

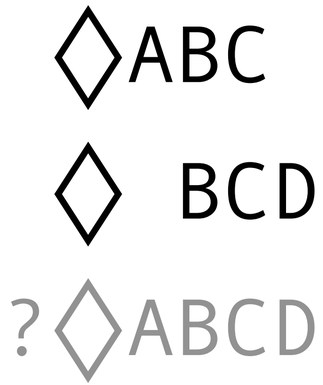
Closed-world belief

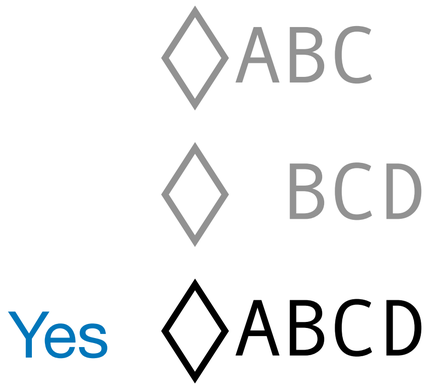


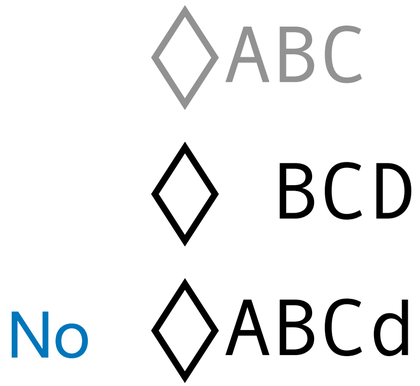
Every state
is covered
by some fact

Combining decision trees

One stronger fact that explains two consistent facts?







Metaphor of learning & memory

Facts acrue monotonically, beliefs are defeasible

Fact trees

- Add new fact
- Strengthen old fact
- Delete two consistent facts with one stronger fact

Swiss cheese

- Add new bubble
- Shrink old bubble
- Delete two overlapping bubbles with one smaller bubble

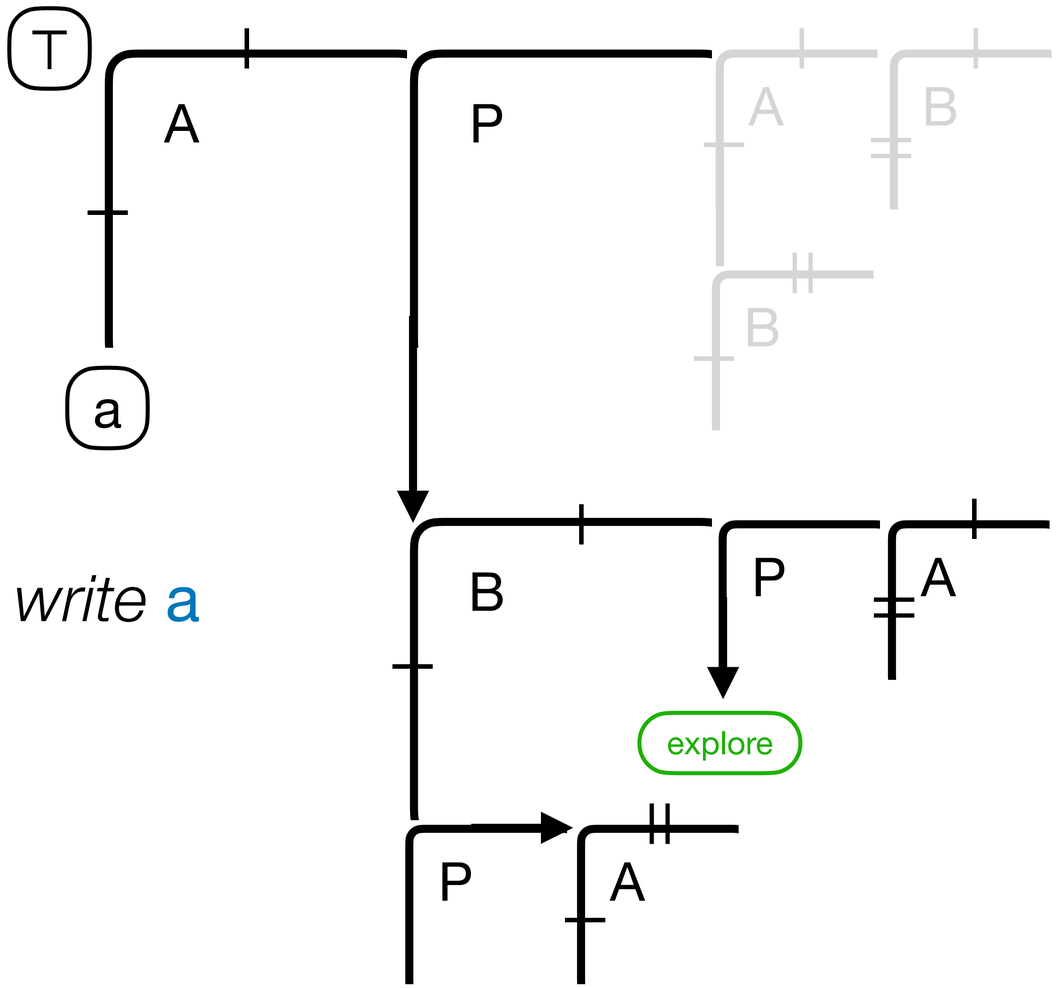
Monotonic increase of distinguished states

- Increase in number of isolated bubbles & decrease in size
- But all these records come at a cost!

Computational complexity of exploration

Kleene automata are learnable by experiment, not observation alone

x	Passive observation	Gold (1978))	Oracle generated examples & counterexamples
✓	Active experiment	Angluin (1987)	Teacher chosen examples (EQs) Student posed queries (MQs) 'Minimally Adequate Teacher'

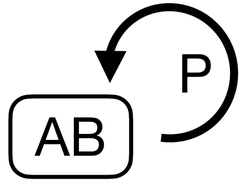


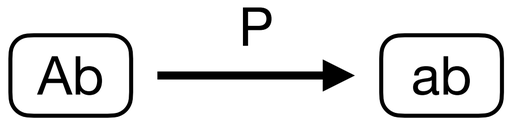
write a

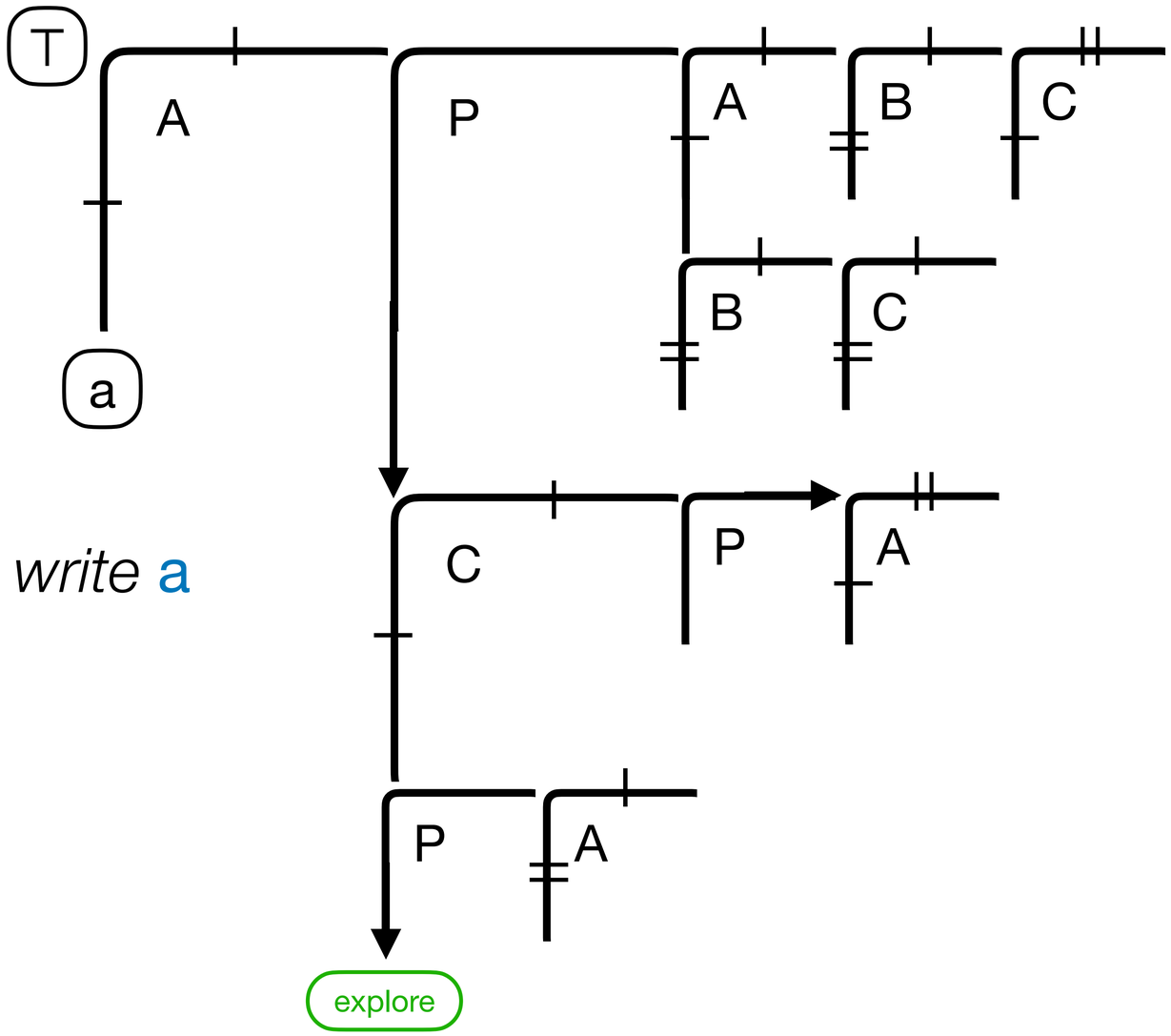
Active experiment

Minimally curious student

APa
APB
APab
ApBP
ApbPa





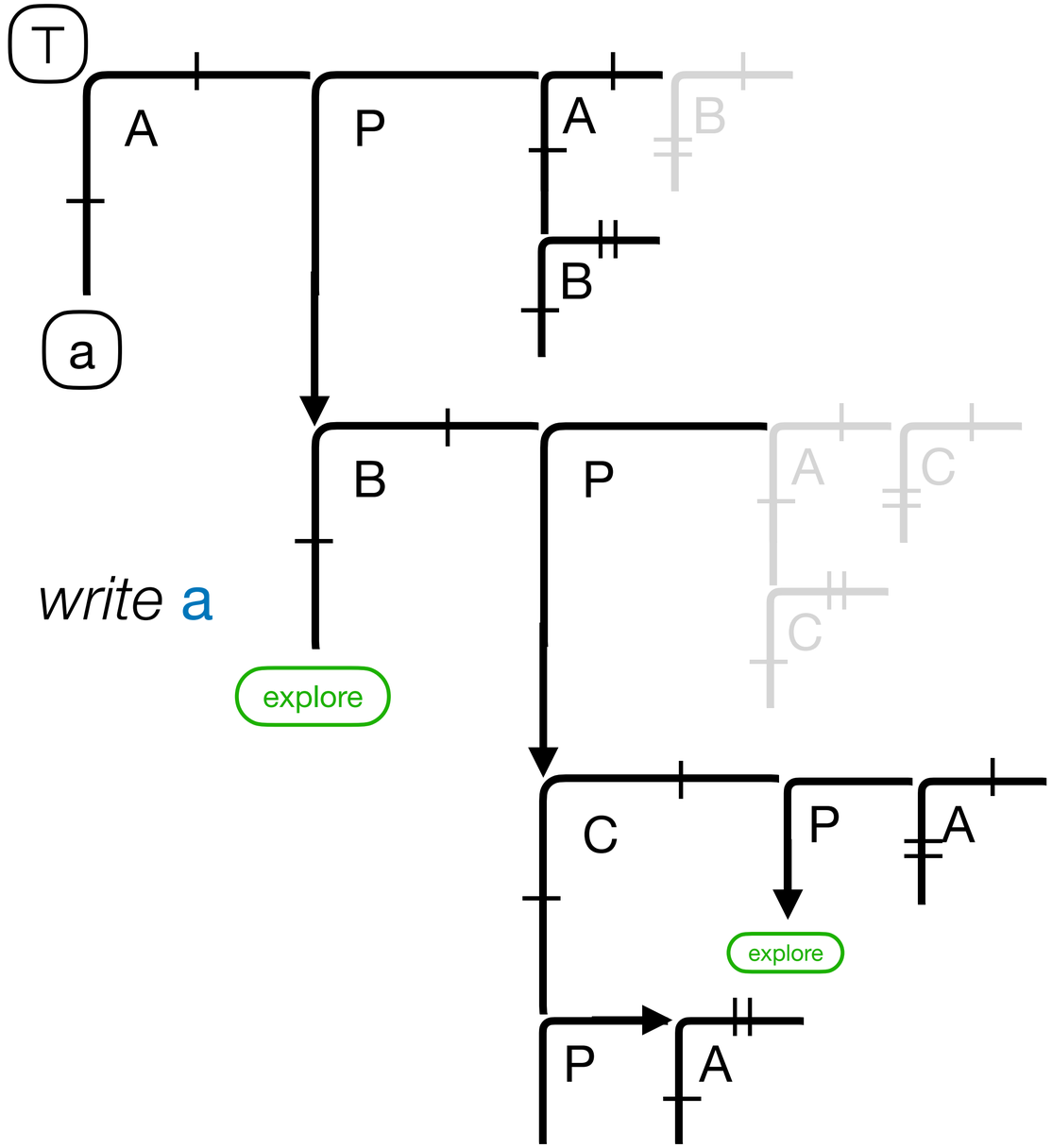


write a

Active experiment

Minimally curious student

APa
APB
APaBC
APABc
ApCPa
ApcPA



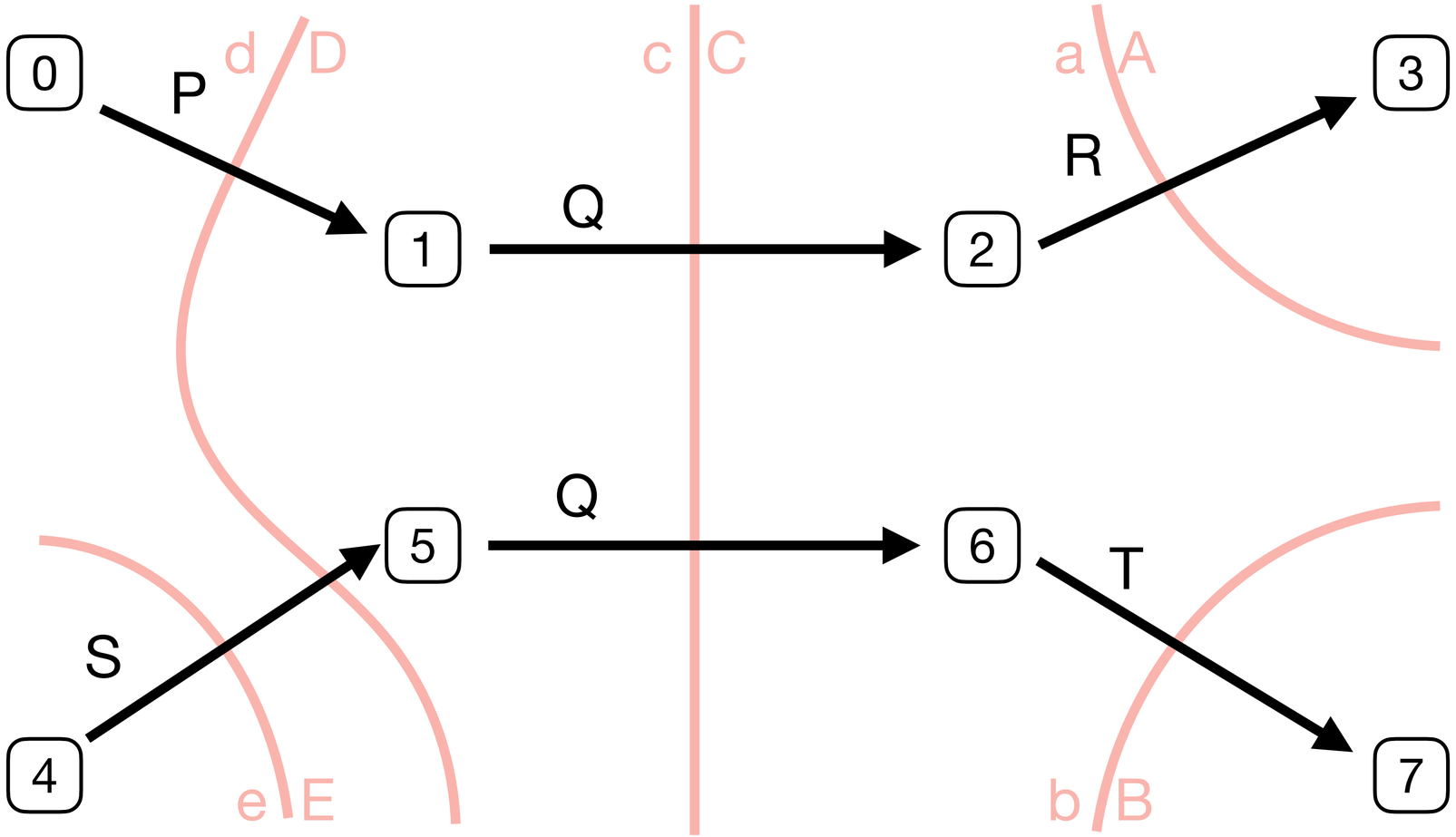
Active experiment

Minimally curious student

APa
APB
APab
ApBP
ApBPac
ApBPC
ApBpcPa
ApBpCP

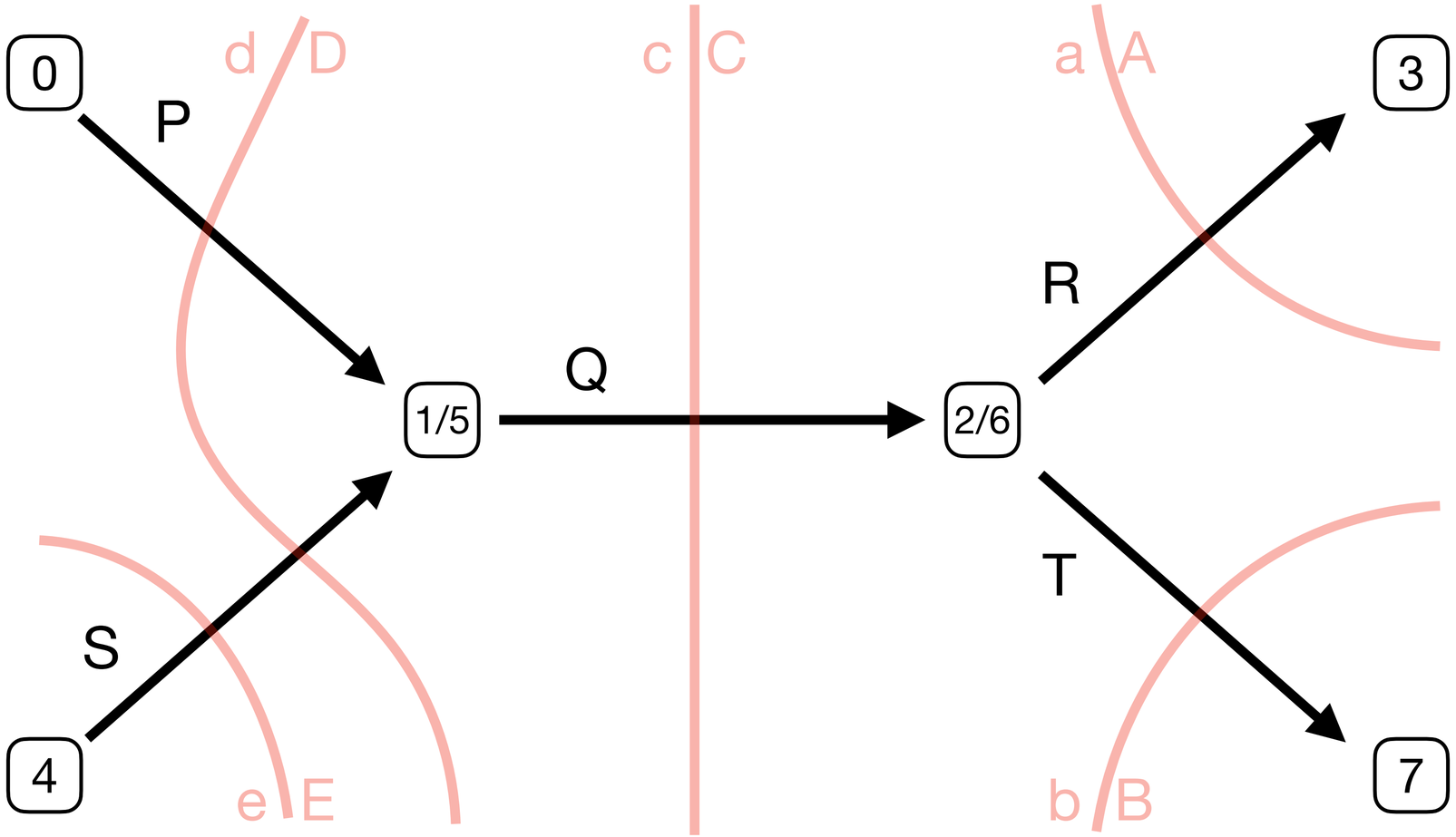
Combining maps

Consistent segments of two maps



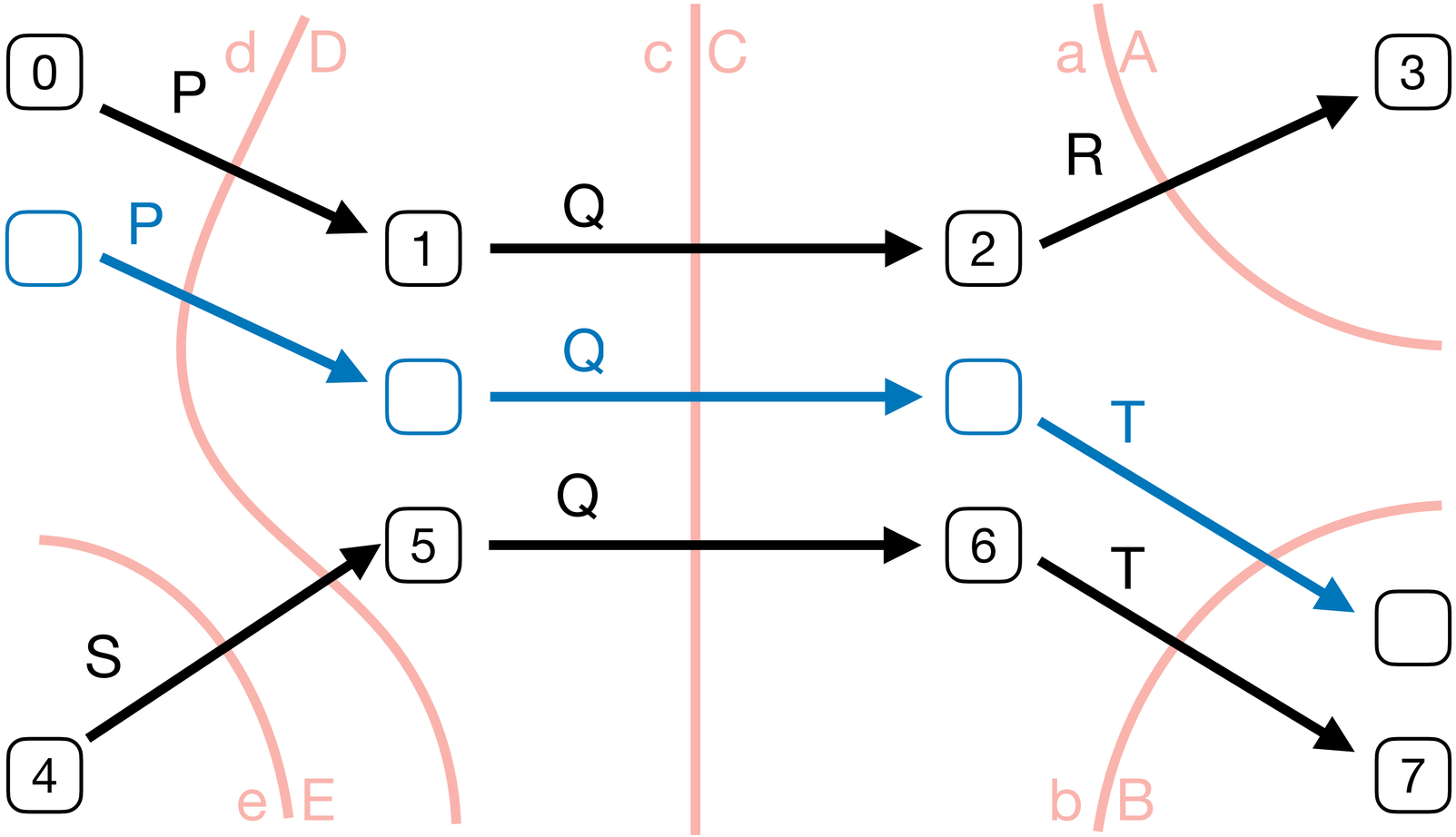
Combining maps

Do these consistent segments reflect one common path in the maze?



Combining maps

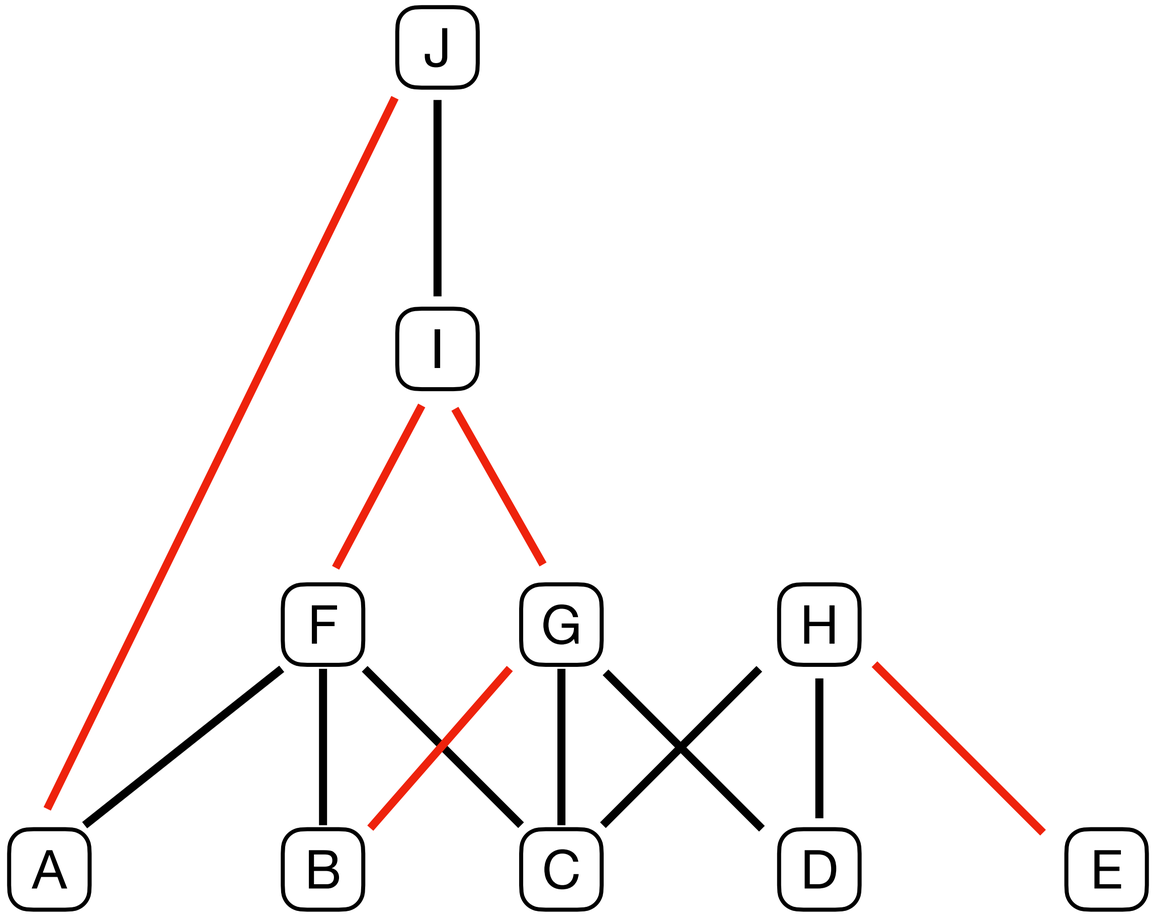
Bisimulation experiment tests for consistent common path

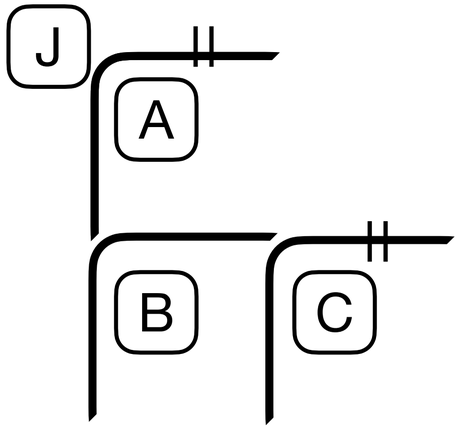


Distribution to lower level tests

Rewriting nested lists to ground tree

```
J = aI
I = fg
H = CDE
G = bCD
F = ABC
E ground
D ground
C ground
B ground
A ground
```





Means-ends hierarchy

viz acting calls on planning at levels below

Neuroscience

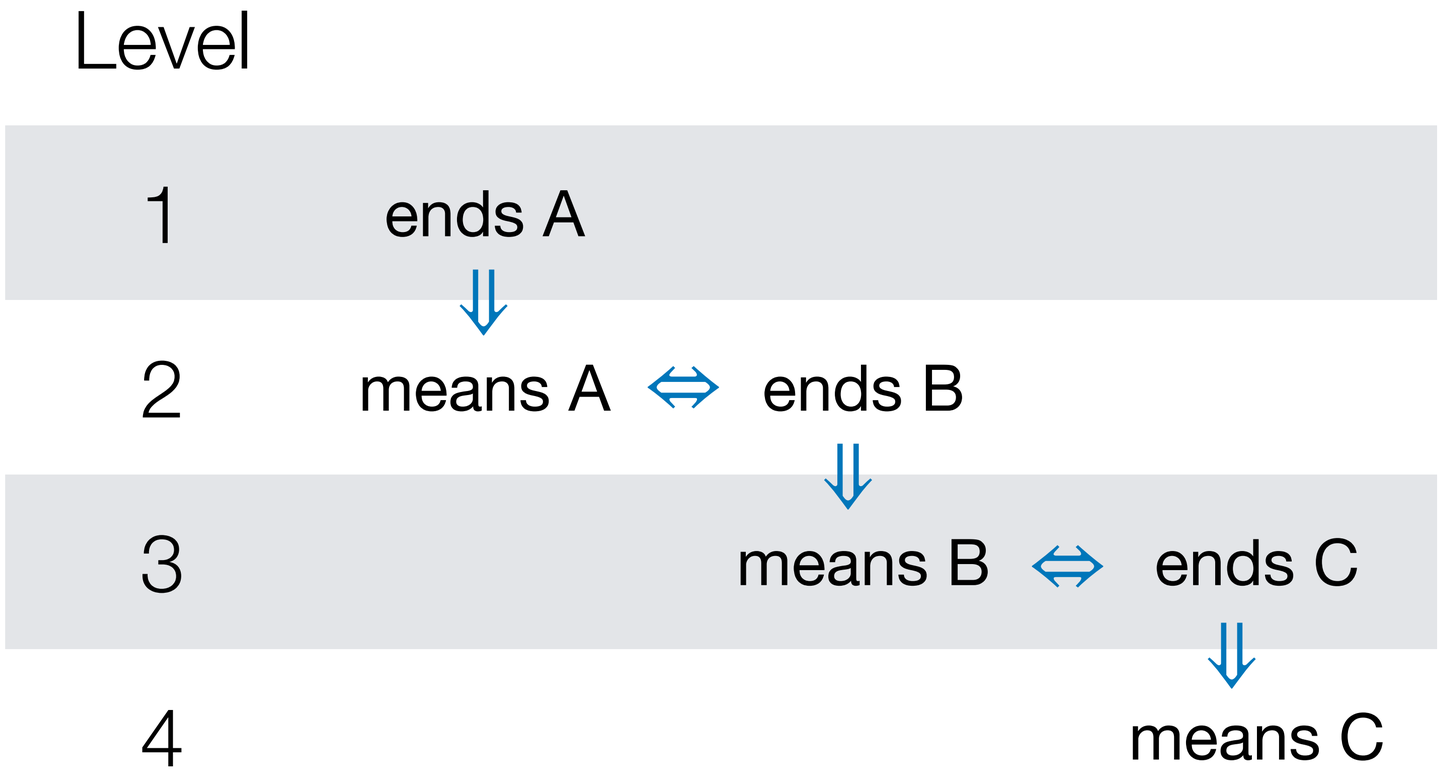
systems consolidation

Computer science

program compilation

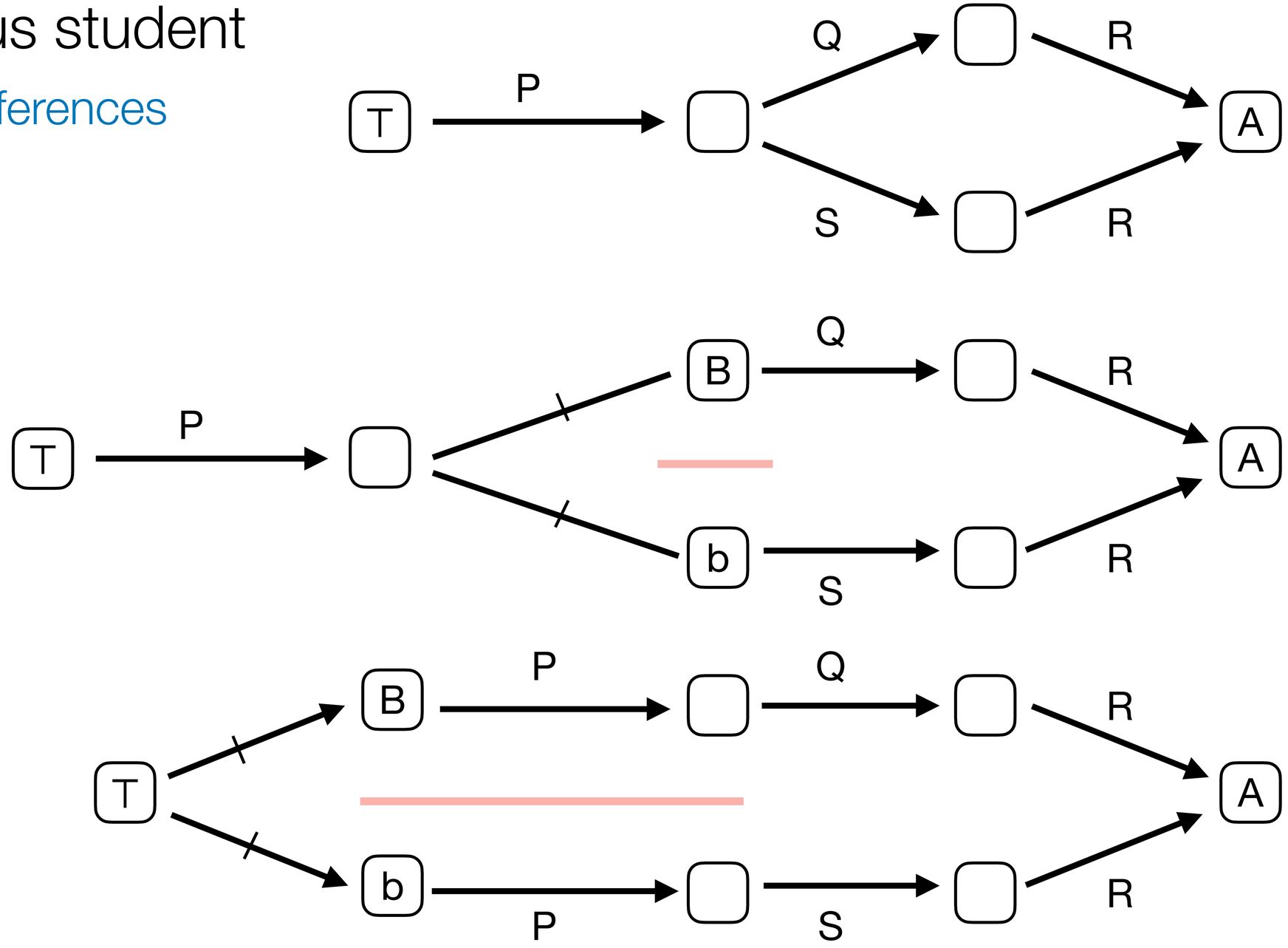
Hierarchical planning

downward distribution



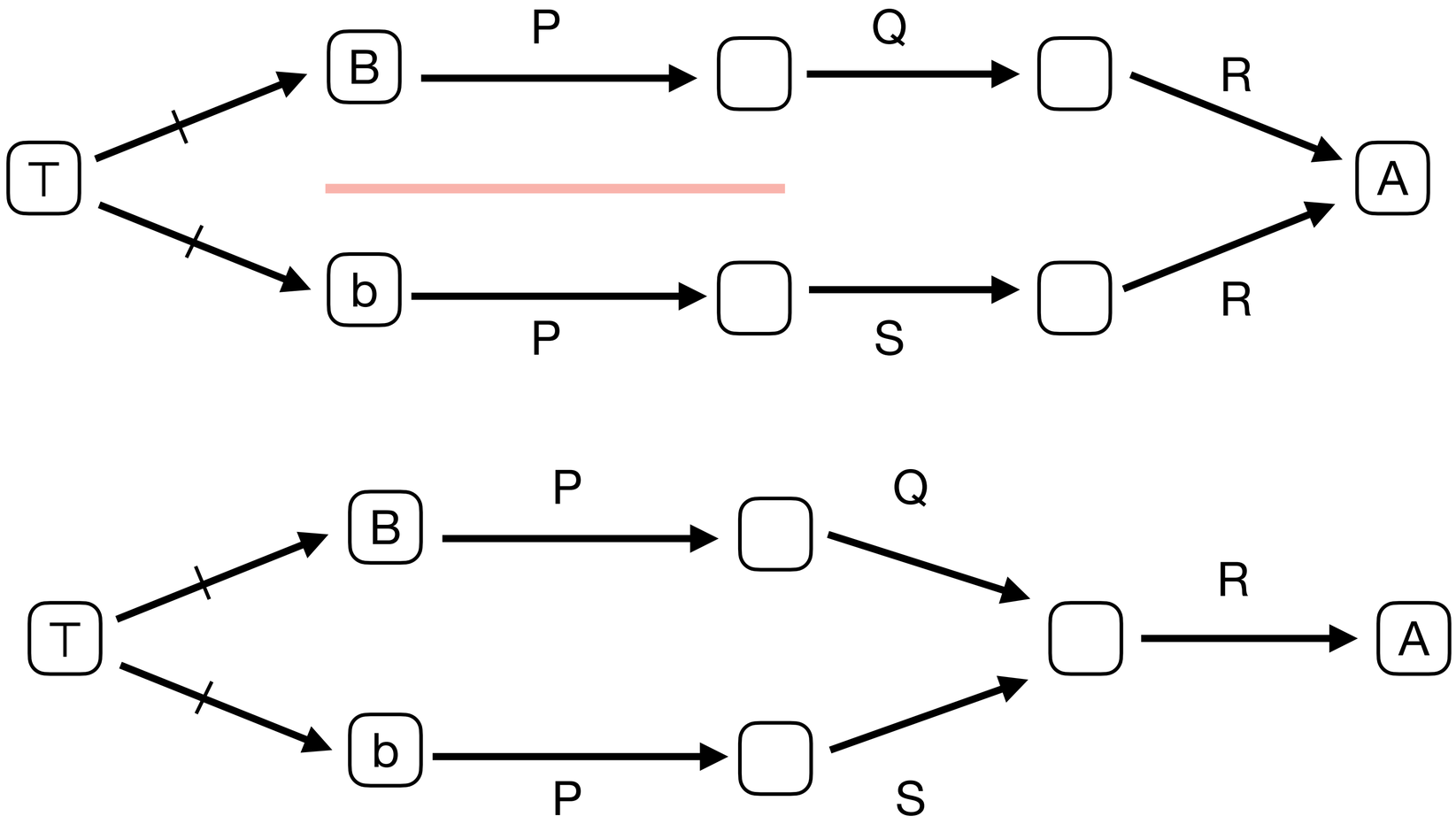
Minimally curious student

Unzips initial differences



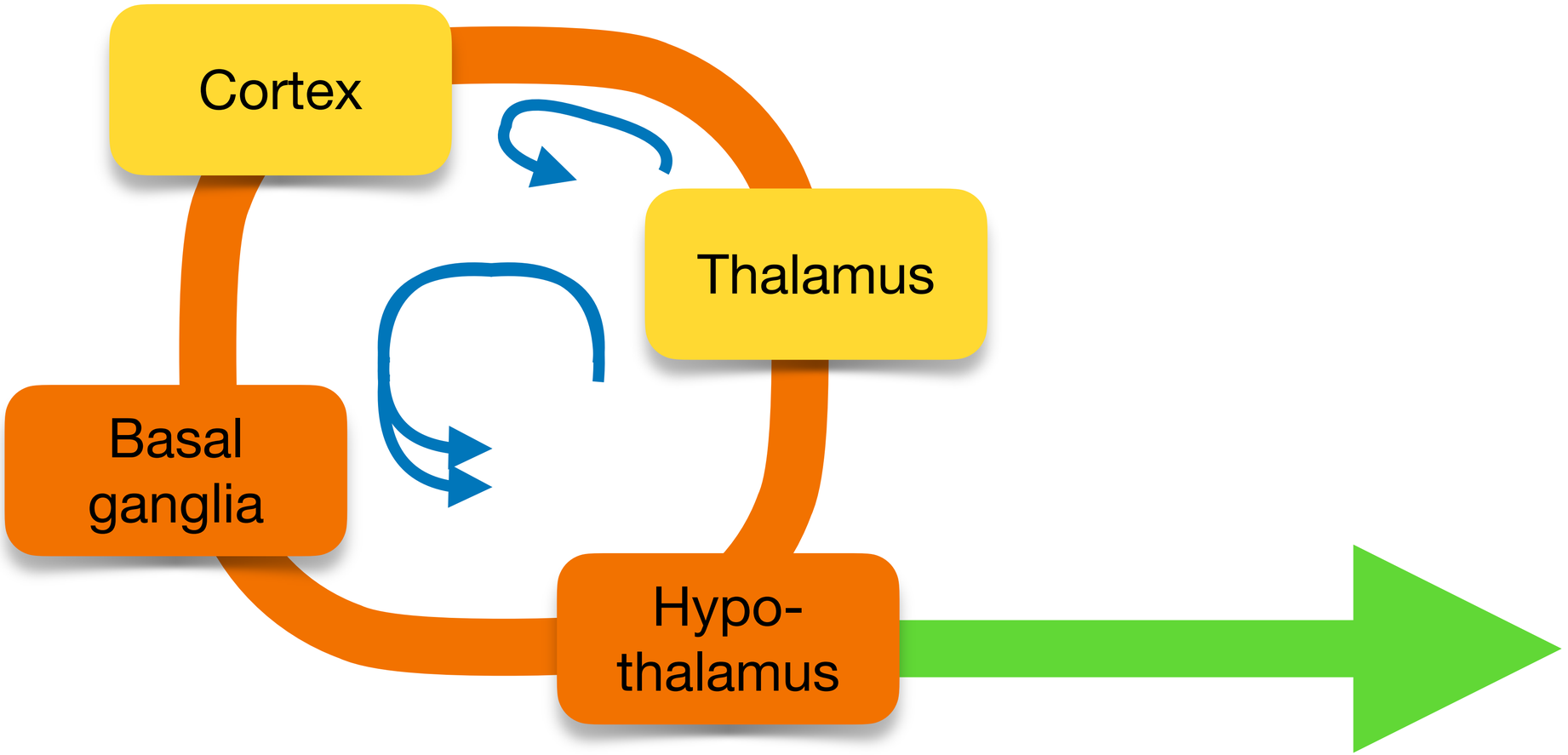
Maximally forgetful professor

Zips final differences



What of the cerebrum?

Formal explanations & neural realizations



Cognitive maps in rats & men

Edward Tolman (1948). *Psychological Review*

Five conditions for narrow maps

- Initial exploration of novel maze
- Feature-poor maze
- Habitual path through a familiar maze
- Extreme urgency or frustration
- Brain injury

Why the amygdala?

Urgency & frustration : shifting between hot & cold

Shift of problems for planning

From neutral to valenced constraints

From appetitive (*write* A) to aversive (*write* a) goals

Shift of mode from planning to acting

From cold reason to hot emotion

- Dr. Banner becomes The Incredible Hulk

From cold to hot games (Conway 1978)

Anytime algorithms & triage

Shift of acting from explore to exploit

Shift of acting from higher-level plans to lower-level habits

Two fundamental cognitive processes

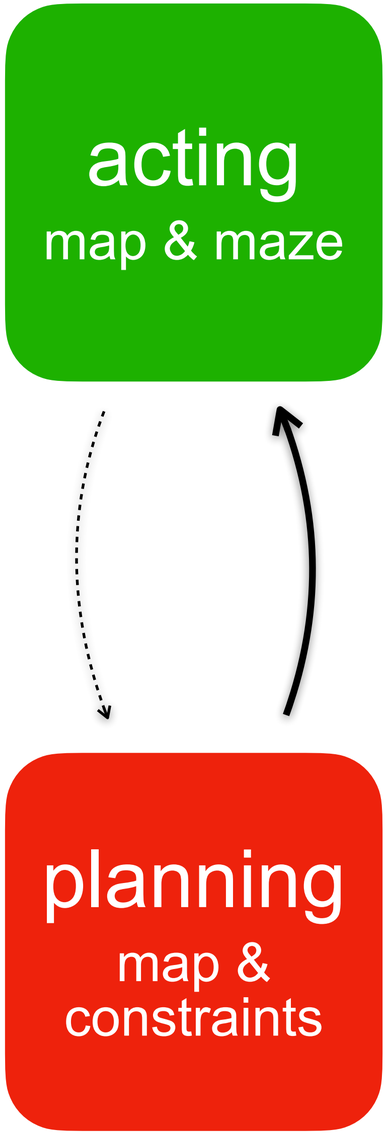
acting & planning across wakefulness & sleep

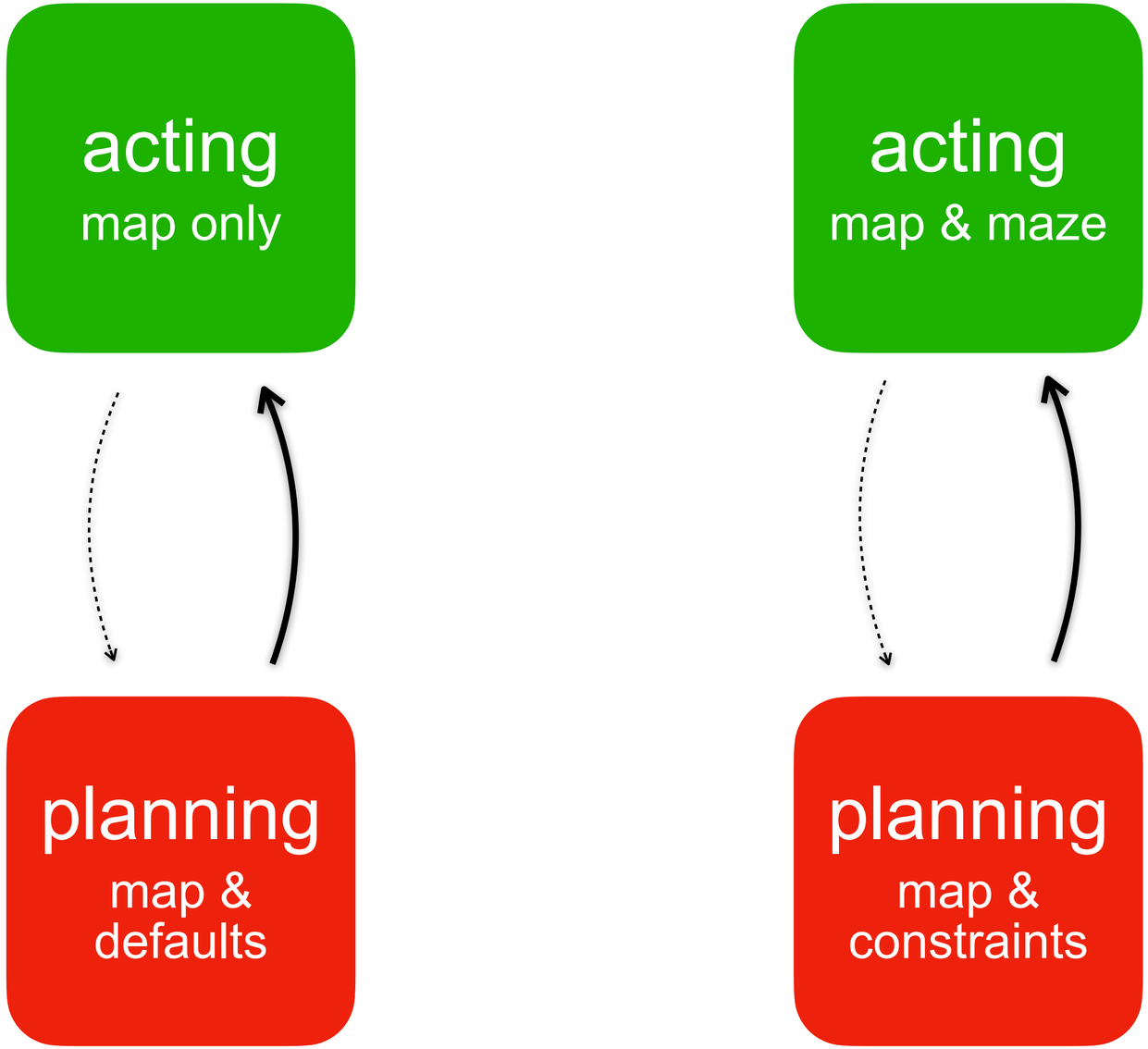
acting : active wakefulness	locomotion
	conscious calculation
	attentive vigilance
offline acting : REM sleep	dreaming
planning : quiet wakefulness	inattentive immobility
	mind-wandering
	N1 drowsy pre-sleep
offline planning : NREM sleep	N2/N3 deep sleep

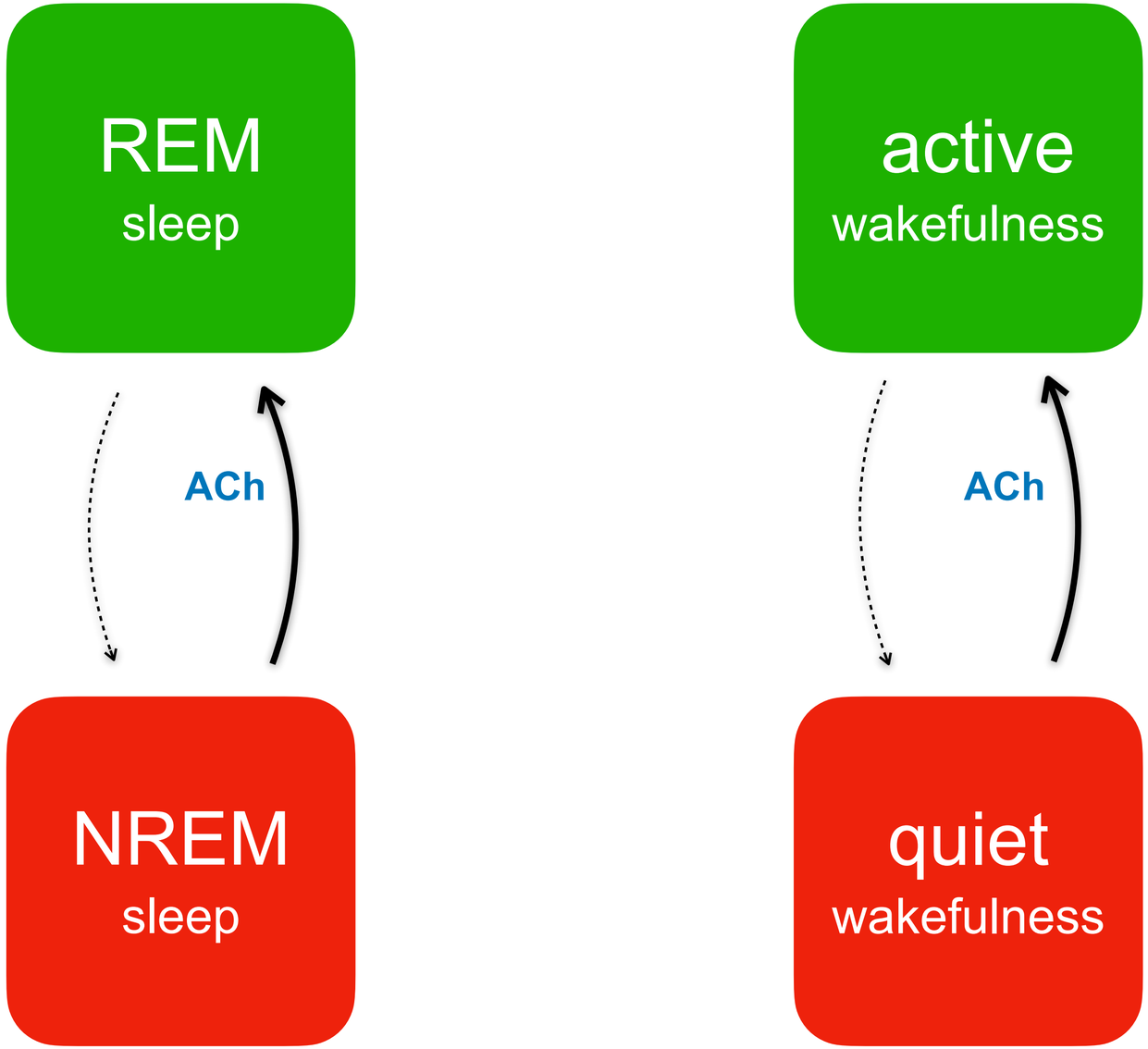
Two fundamental brain modes

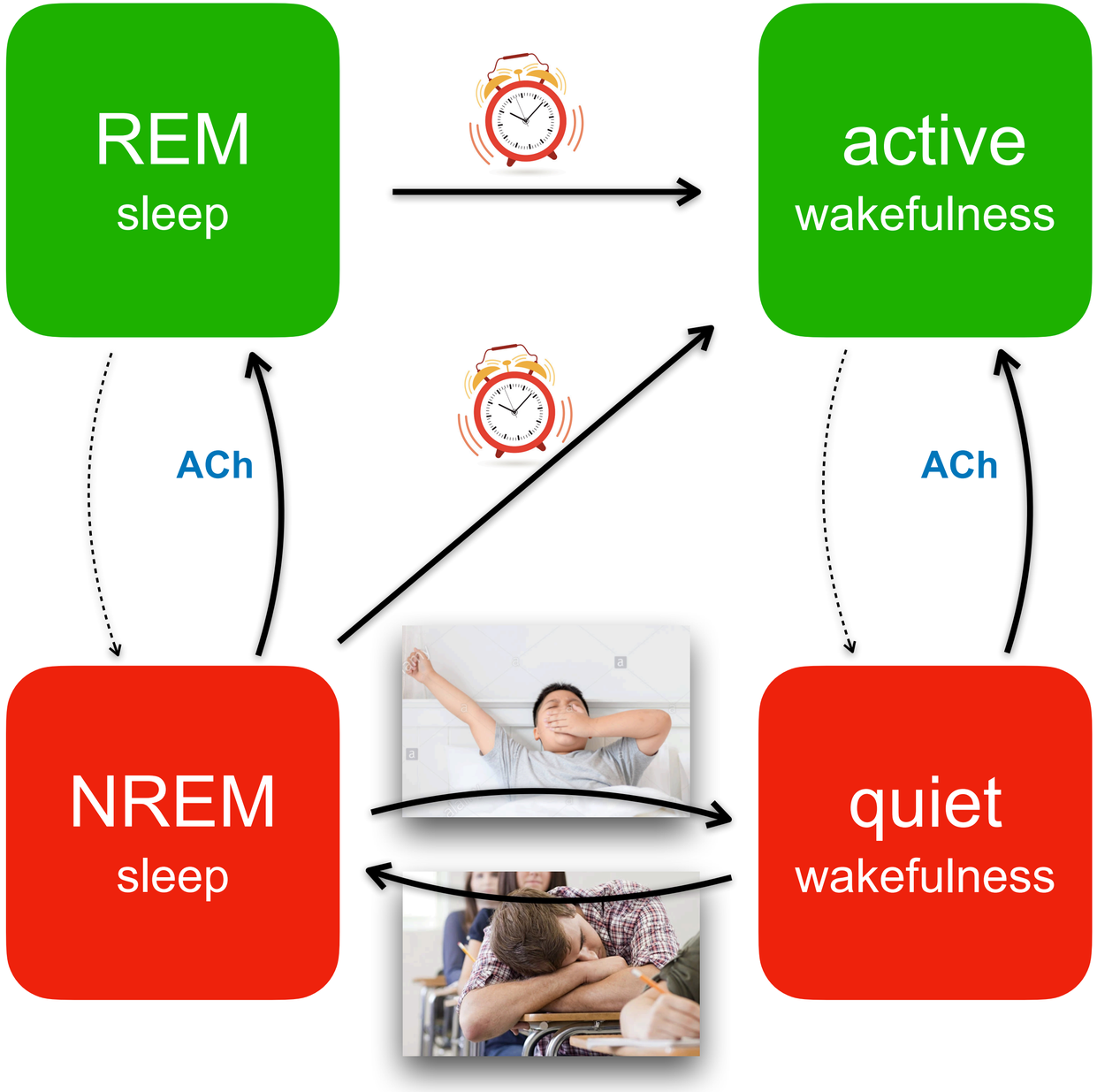
acting & planning across wakefulness & sleep

	hippocampus	cortex	cortico-thalamic	skeletal muscles	eye muscles
active	theta 4-8 Hz	gamma 30-90 Hz		active	saccades
REM	theta 4-8 Hz	gamma 30-90 Hz		silenced	REMs
quiet	SWR 100-200 Hz	slow / delta < 4 Hz	alpha 9-15 Hz	quiet	quiet
NREM	SWR 100-200 Hz	slow / delta < 4 Hz	spindles 9-15 Hz	quiet	quiet


acting
map & maze
planning
map &
constraints

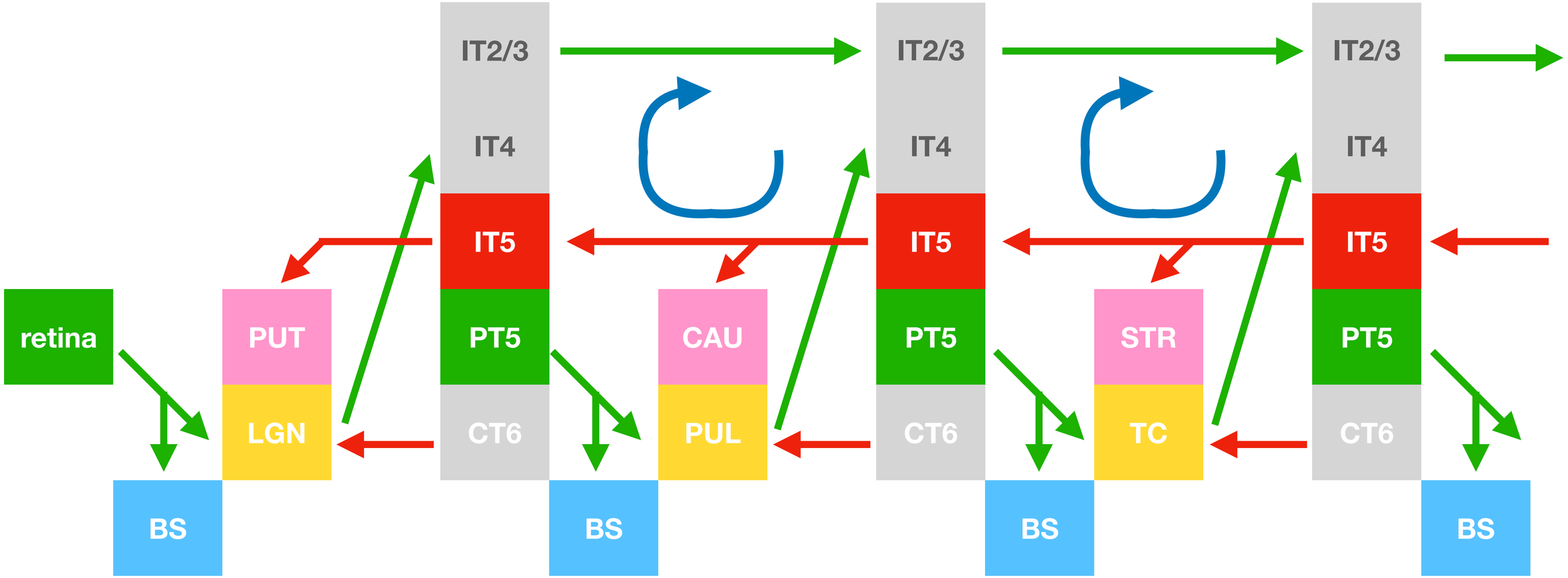

acting
map only
planning
map &
defaults
acting
map & maze
planning
map &
constraints


REM
sleep
ACh
NREM
sleep
active
wakefulness
ACh
quiet
wakefulness


REM
sleep
active
wakefulness
ACh
ACh
NREM
sleep
quiet
wakefulness

Reciprocal paths in the cerebral map

forward (bottom-up) / backward (top-down)



Planning, acting & learning

Neural phenomena? Neural realizations? Of maps, plans & rewrite rules

Planning with map of past runs & current goal

Current goal specified as *writes* & *stays*

- attentional modes

Backward-step from winning positions on the map

Search for satisfying forward-paths & mark these routes

If necessary, test the prestate into cases & mark both alternatve routes

Acting as following the plan

Run forward-path marked on the map

- theta alternation : test · move · test · move · test
- frame assumption, theta precession : once a feature is tested, value attended in subsequent run

Passive learning from counterexamples

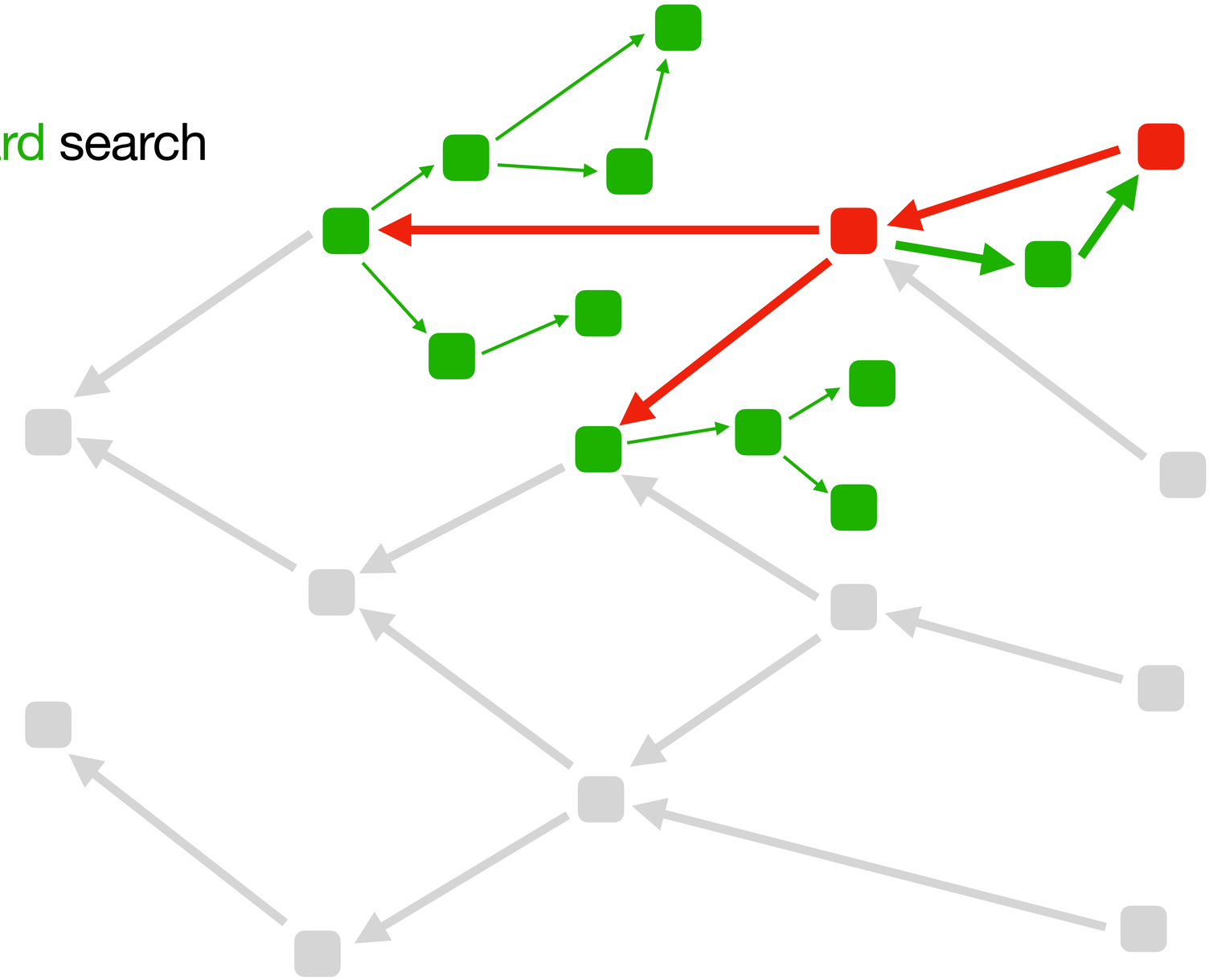
- Update map wherever run breaches plan (CWA violations)

Active learning from experiment

- Unplanned exploration (default biases for sampling moves & tests)
- Simple experiment (MCS)
- Bisimulation experiment (MFP)

Basic planning

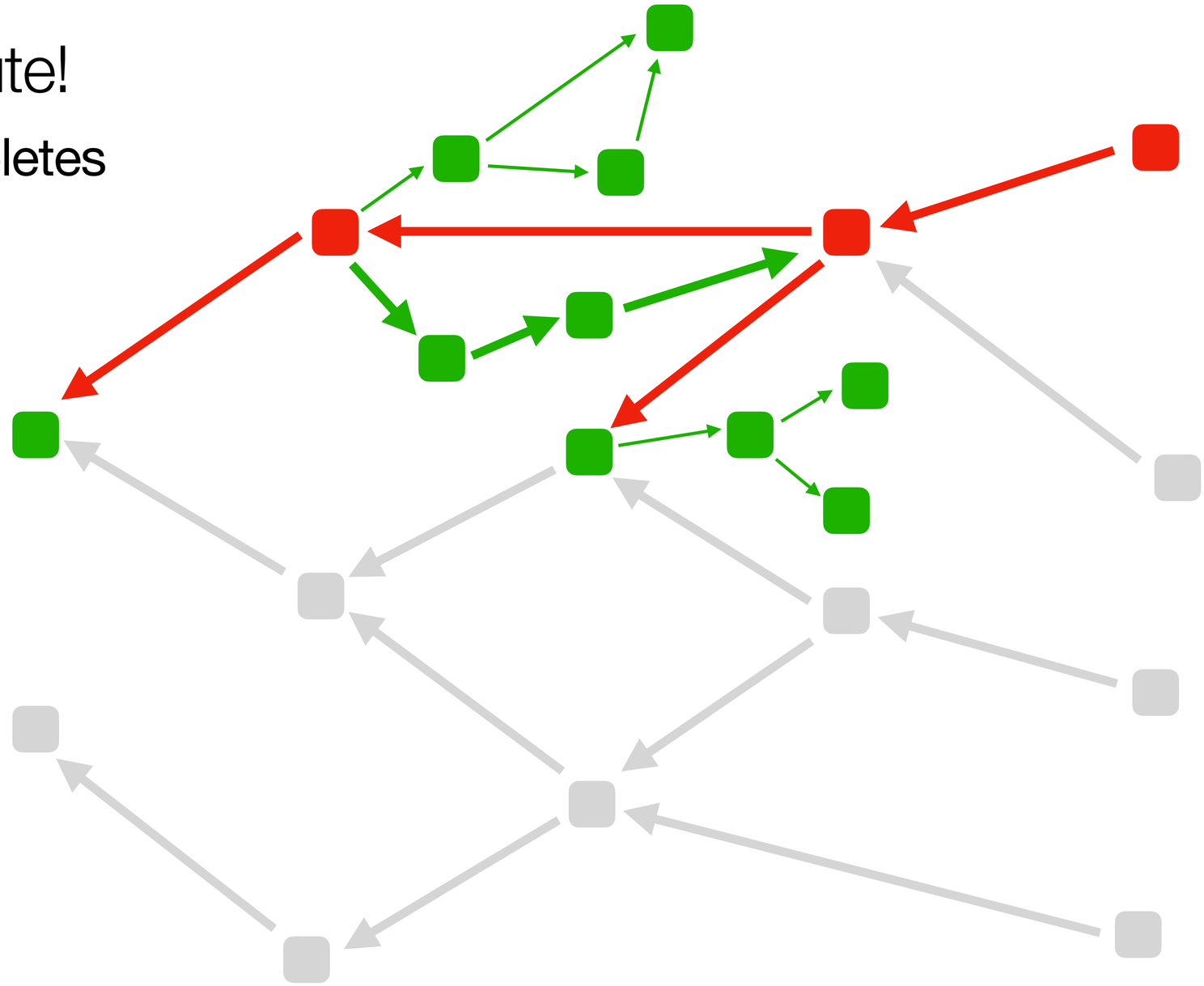
backward-then-forward search



Marking intended route!

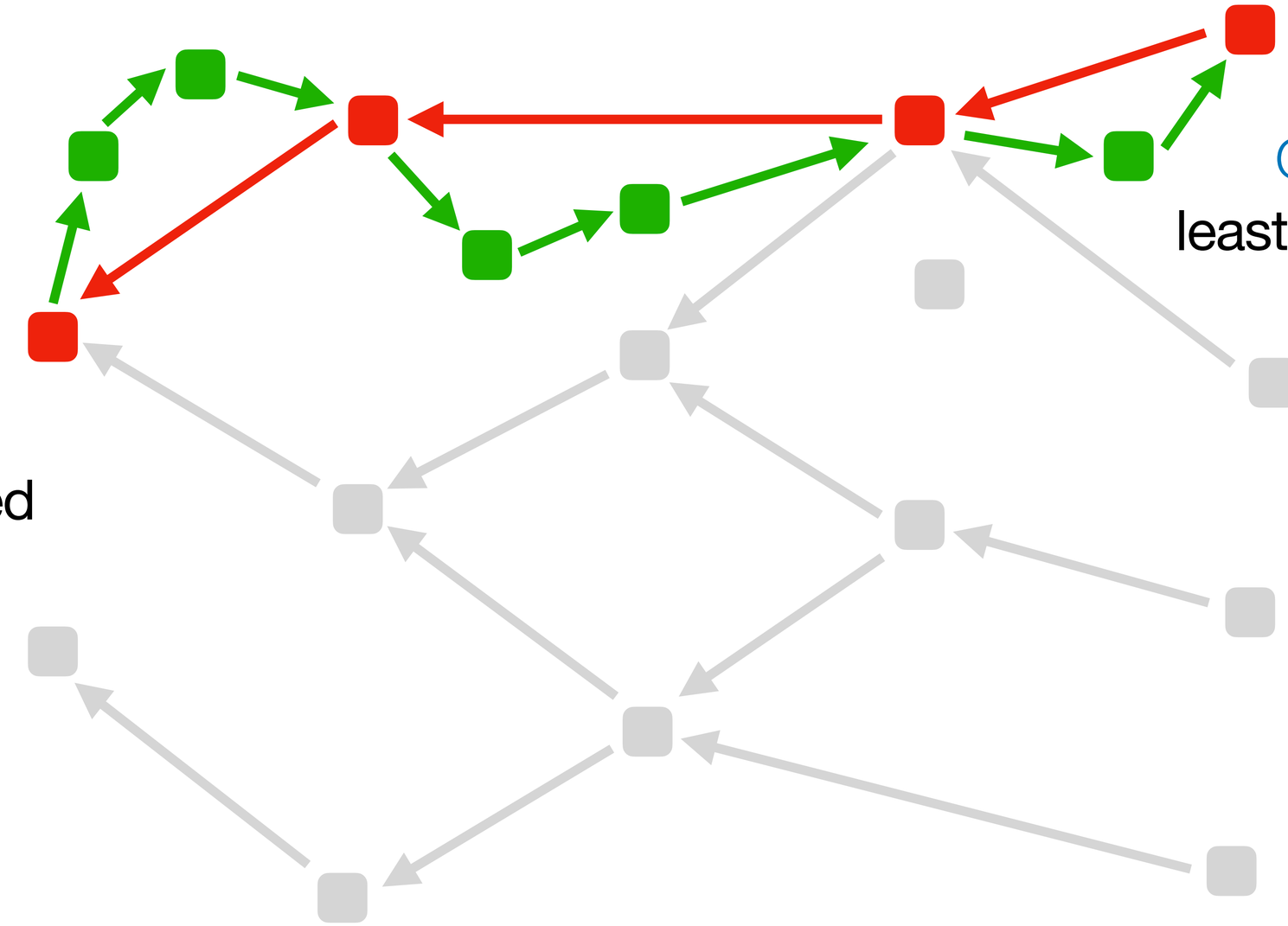
forward path that completes

backward step



Acting

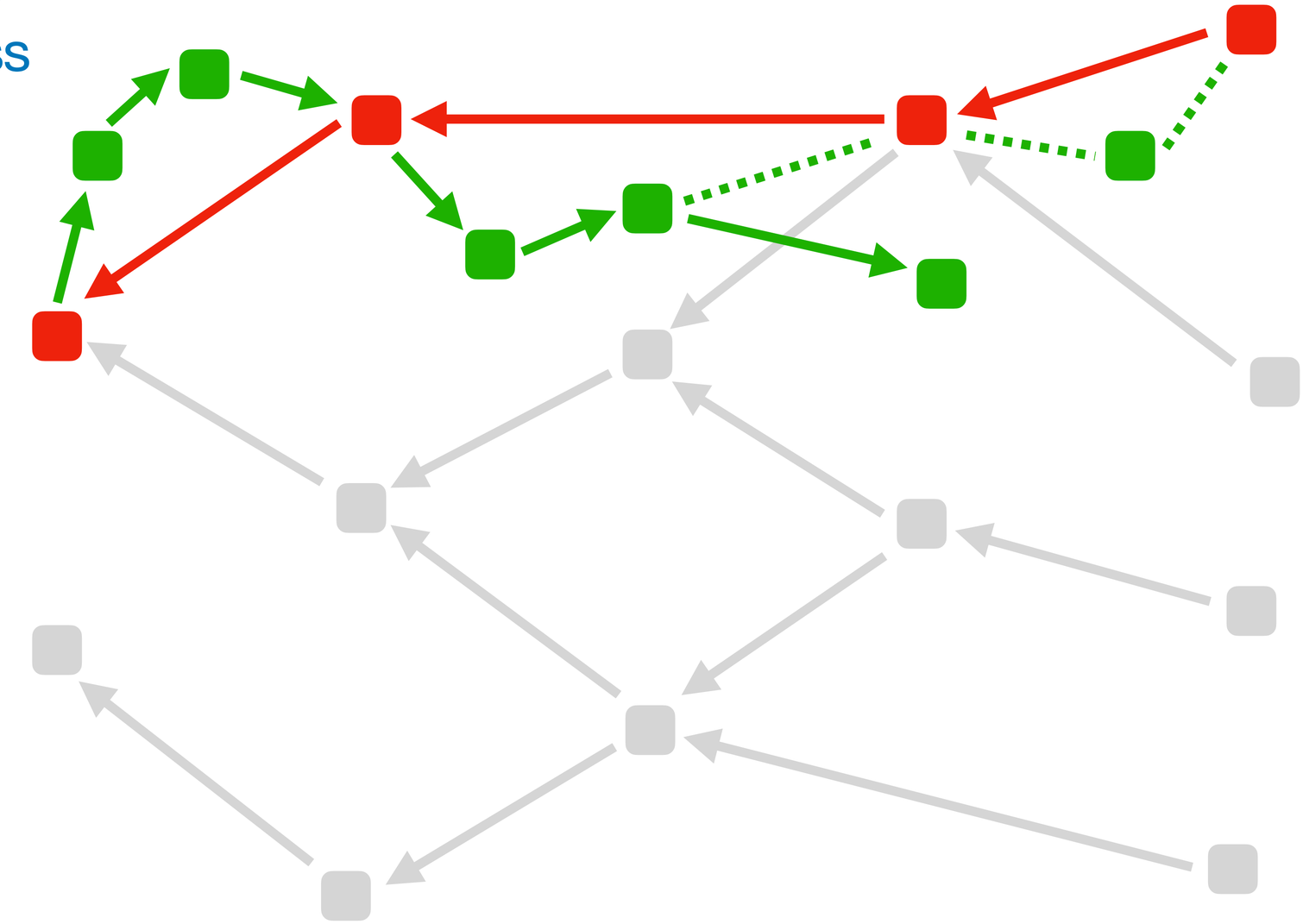
Completed plan is good to go!



Learning from experience

from unexpected failure

from unplanned success



What of natural language?

Parsing as planning

Natural language

- Bisimulation experiment within the same? Or in two different crania?
- Sentence is one run through cognitive map
- Discourse is a larger map fragment

Downward backward-then-forward planning

- higher-level step ≻ lower-level path
- rewrite context-free grammar

Upward forward-then-backward parsing

- lower-level path ≻ higher-level step
- rewrite left-corner parser

$$Q \rightarrow RST$$

$$U \rightarrow WXYZ$$

PQUV

PRSTUV

PRSTWXYZV

Acknowledgements

Advisors & teachers

Sydney Brenner, Bozena Dostert, James Olds, Richard Russell, John Sulston, John White

Colleagues

Maurice Bessman, Harshida Bhatt, Joe Culotti, Andy Fire, Dave Hall, Phil Hartman, Carl Johnson, Tony Otsuka, Jim Rand, Aaron Shatkin, Andrew Spence, Nichol Thomson, Haiqing Zhao

Postdocs & students

Matthew Buchner, Chaobo Guo, Harald Hutter, Naoaki Ishii, Keiichiro Joh, Edward Kipreos, Surjeet Mastwal, Carolyn Norris, John Plenefisch, Brian Stern, Bill Wadsworth

Co-author

Rui Proença

Host

Scott Emmons

Distance in maze? In map?

Metrics for moves or tests

Metric axioms

$i\,i = 0 \quad i\,j = j\,i \quad i\,k \leq i\,j + j\,k$

Shortest path (egocentric)

length of shortest known path : decreases as we find shortcuts

Boolean difference (allocentric)

sum of known differences : increases as we find differences

A* algorithm, or best-first graph search (Hart et al 1968)

cost_estimate = moves_taken + differences_remaining

admissible heuristic if we use (at most) 1-Hamming moves

Few events have direct converses

You can't just run a plan backward!

Locomotion seems to have direct converses

forward ∘ backward = backward ∘ foreward = *no-op*

northward ∘ southward = southward ∘ northward = *no-op*

Throwing a ball

Catching a ball is not its converse.

Whatever is unthrowing a ball? A yo-yo?

Eating an apple

If hungry & apple nearby then

reach-to-grasp; hand-to-mouth;

while not-done (bite; chew; swallow);

end.

Vomiting is not its converse.

Tripping on a rug

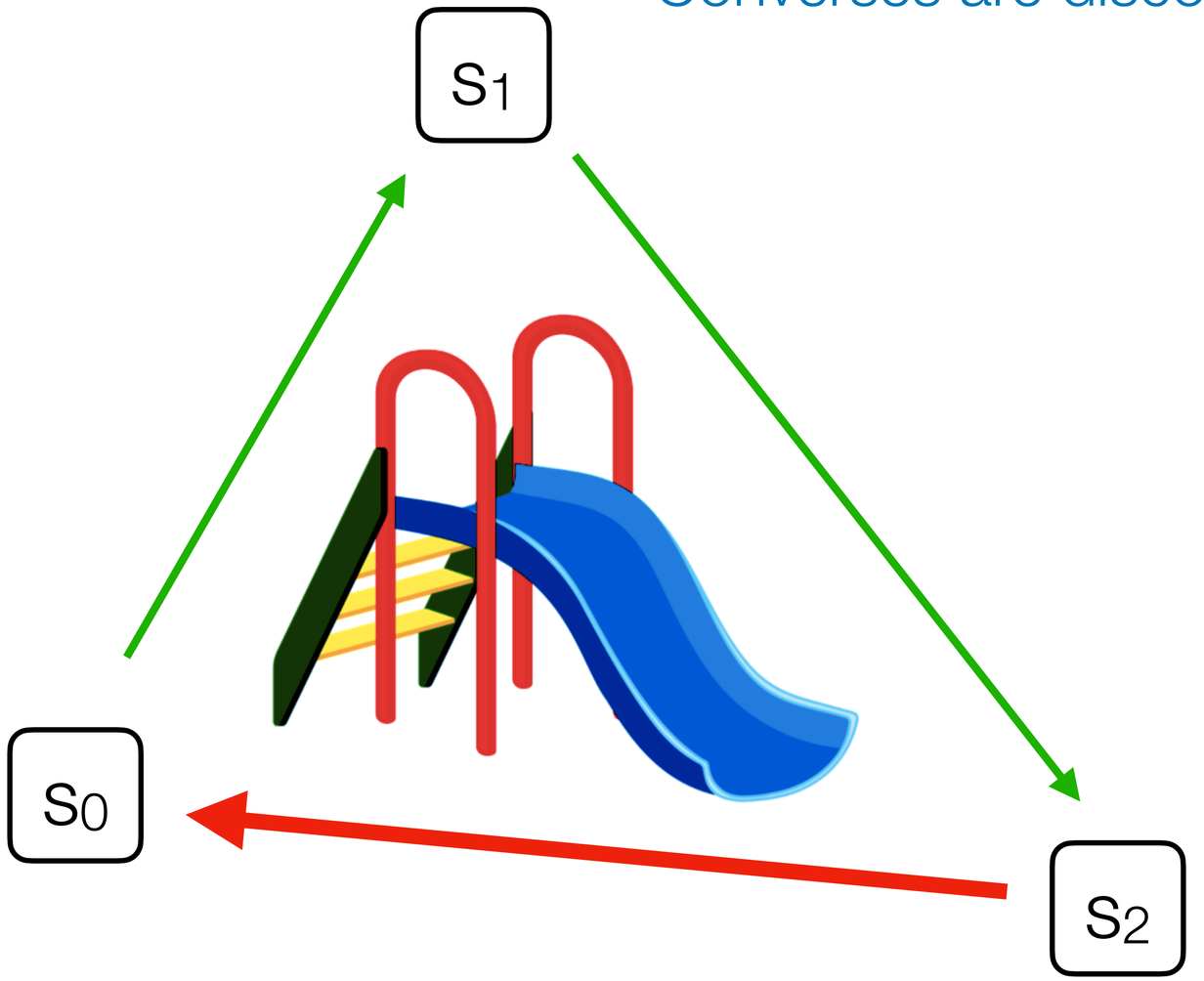
Learned converse : recovering ones balance

Dropping a glass of water

Learned converse : clean-up spill; pour another.

Practical converses are learned loops

Converses are discovered, not innate!



Loop a playground slide

(climb-up the ladder; slide-down the chute)* ad nauseum

Retracing a path

Path converse from step converses

Forth & back through a blind alley

Follow path; turn-around; retrace steps

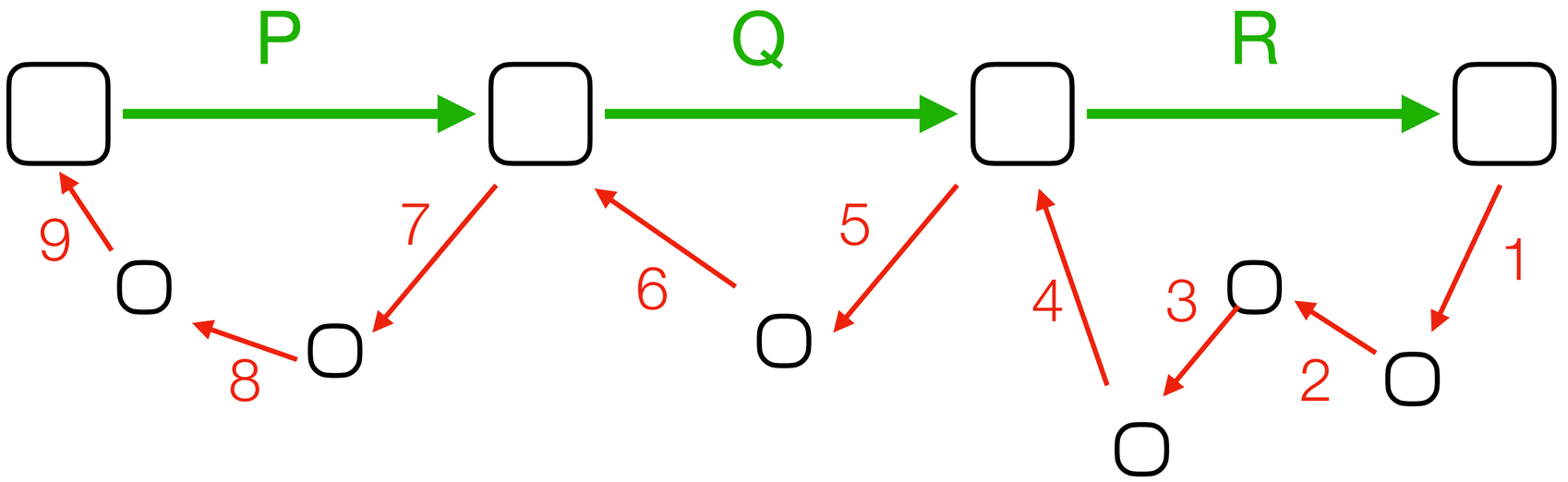
start $P \circ Q \circ R \circ R^{-1} \circ Q^{-1} \circ P^{-1}$ finish

start ~ finish

Path converse is reversed path of step converses

$(P \circ Q \circ R)^{-1} = R^{-1} \circ Q^{-1} \circ P^{-1}$

Sequences of moves & their converses



$(p \circ q \circ r)^{-1} = r^{-1} \circ q^{-1} \circ p^{-1}$

$= (1 \circ 2 \circ 3 \circ 4) \circ (5 \circ 6) \circ (7 \circ 8 \circ 9)$

$= 1 \circ 2 \circ 3 \circ 4 \circ 5 \circ 6 \circ 7 \circ 8 \circ 9$

$(p\ q\ r)(p\ q\ r)^{-1} = p\ q\ r\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9 =$ *no-op*

Loops in the maze? In the map? In the plan? In planning?

H

Alternative paths from si to sj

Two forward paths with the same start and same finish, alternative paths as in Hampton Court maze; not a loop in directed graph

Iteration in plan

- One forward path that meets itself u.vK ...iteration, while program, loop in directed graph
- loop can be divided in two segments, each converse of other

Paths & converse in planning

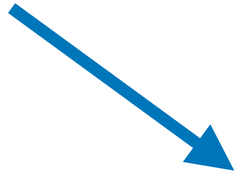
- path & converse loop aka backward then forward, or foward then backwar
- loop divided into two segment, one forwad, one backward

Loops in the maze? In the map? In the plan? In planning?

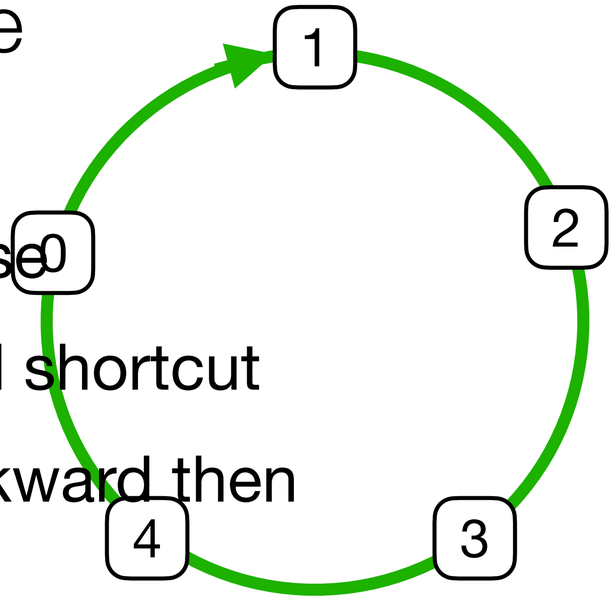
H

1 aBc

Q



Loop a playground slide



forward loop is path.converse

alternatve paths or path and shortcut

loop in the map only of backward then forward

Enumeration of sentences/paths

Generative rewrite algorithms

Queue initialized with non-terminal S
Scan first element of queue for matches to rewrite rule LHS
For each match replace with RHS and add to back of queue
If original has no matches & no non-terminals keep sentence
If no matches & has terminals discard blind alley
Repeat

Generative versus parsing algorithms
Exploration = generative
pathsofar.move.tests(
(0-Hamming add next post-test or try next move
Exploitation = parsing

Context-free grammars

Left corner parsing

non-terminal symbols {S, T, U, V, W, X, Y, Z}

terminal symbols {K, L, M, N, O, P, Q, R}

start configuration [s | terminal sentence]

Chomsky productions

`W → XY | P`

Griebach productions

`W → PXYZ`

Distance in maze? In map?

Metrics for moves or tests

Metric axioms

$i\,i = 0 \quad i\,j = j\,i \quad i\,k \leq i\,j + j\,k$

Shortest path (egocentric)

length of shortest known path : decreases as we find shortcuts

Boolean difference (allocentric)

sum of known differences : increases as we find differences

A* algorithm, or best-first graph search (Hart et al 1968)

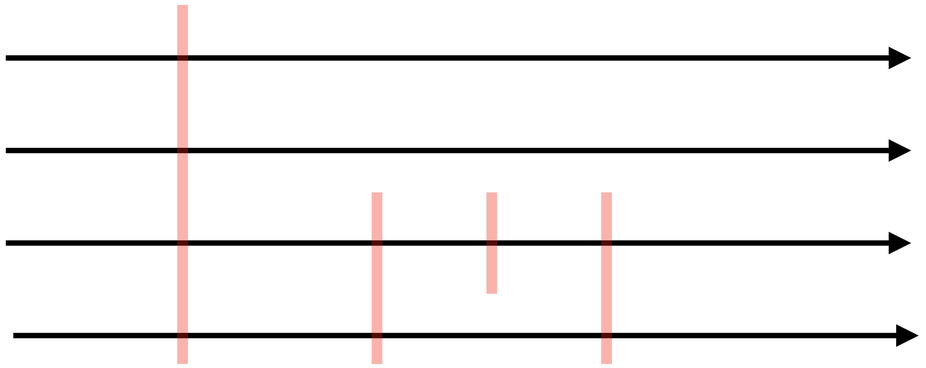
cost_estimate = moves_taken + differences_remaining

admissible heuristic if we use (at most) 1-Hamming moves

Segmentation of experience

The pasta maker

time ⟶



How are runs segmen

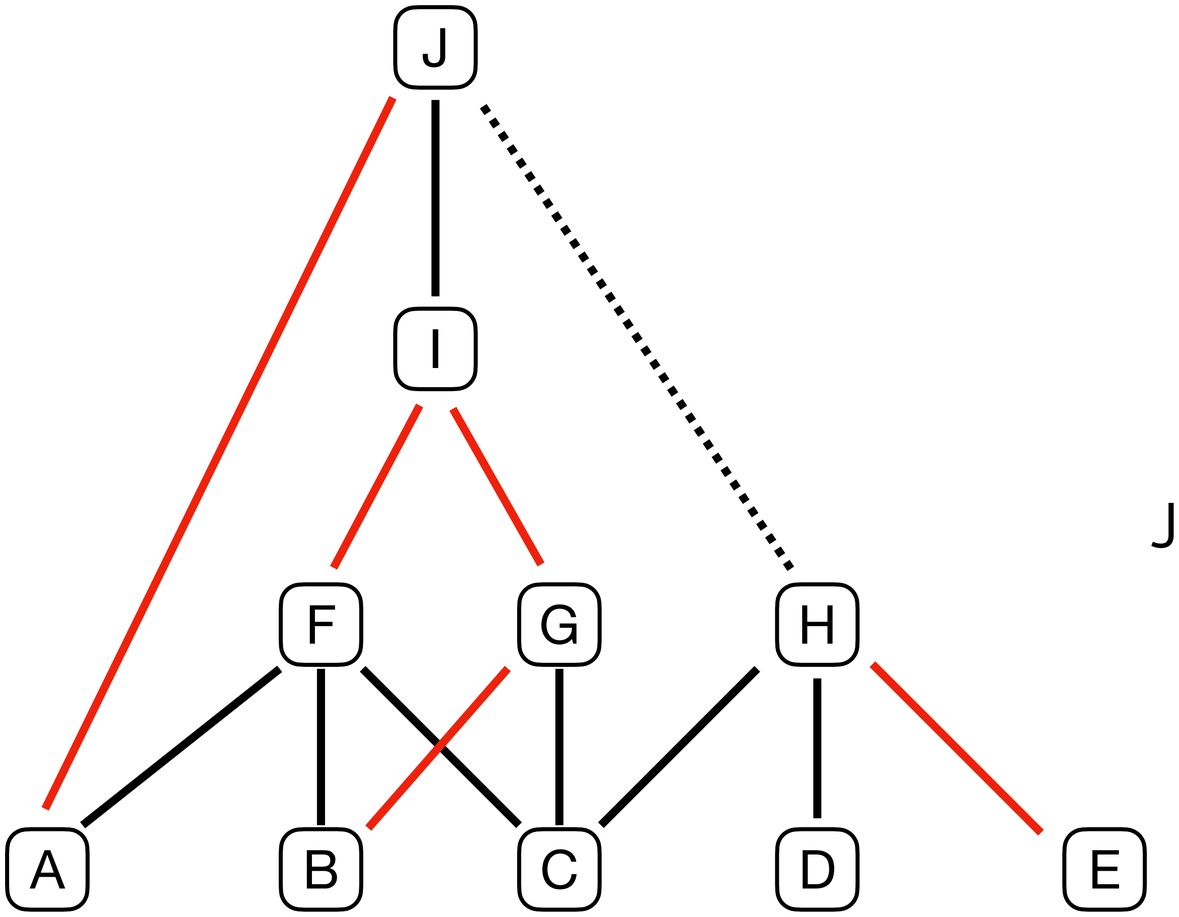
Any Boolean test segm

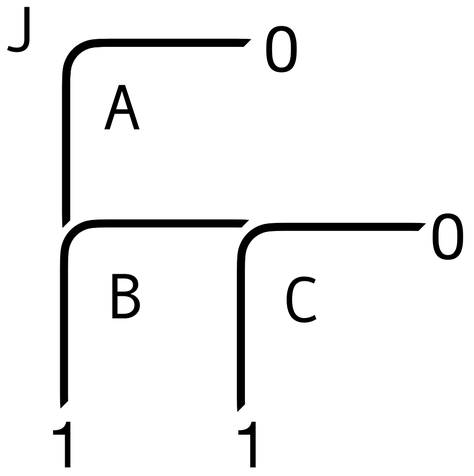
How many raw traces

IDEA each primary ha

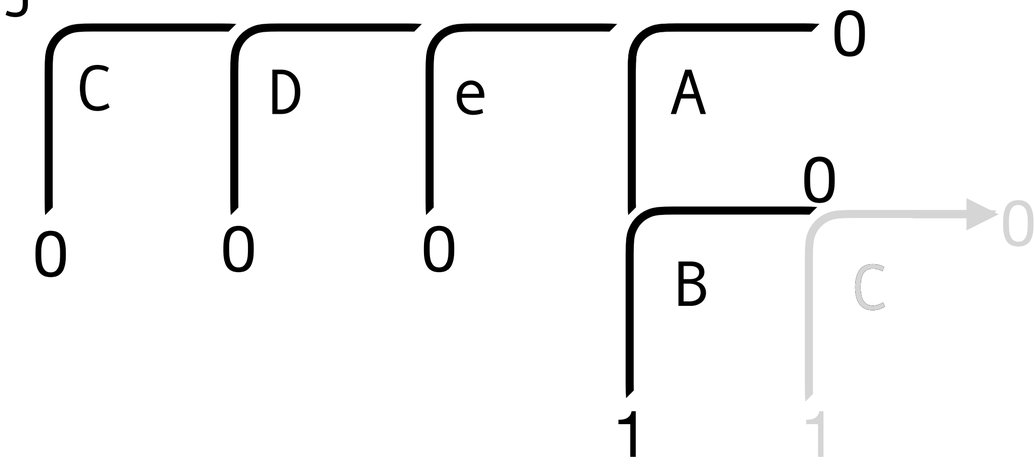
Consolidation to lower level tests

Rewriting nested lists to flat tree





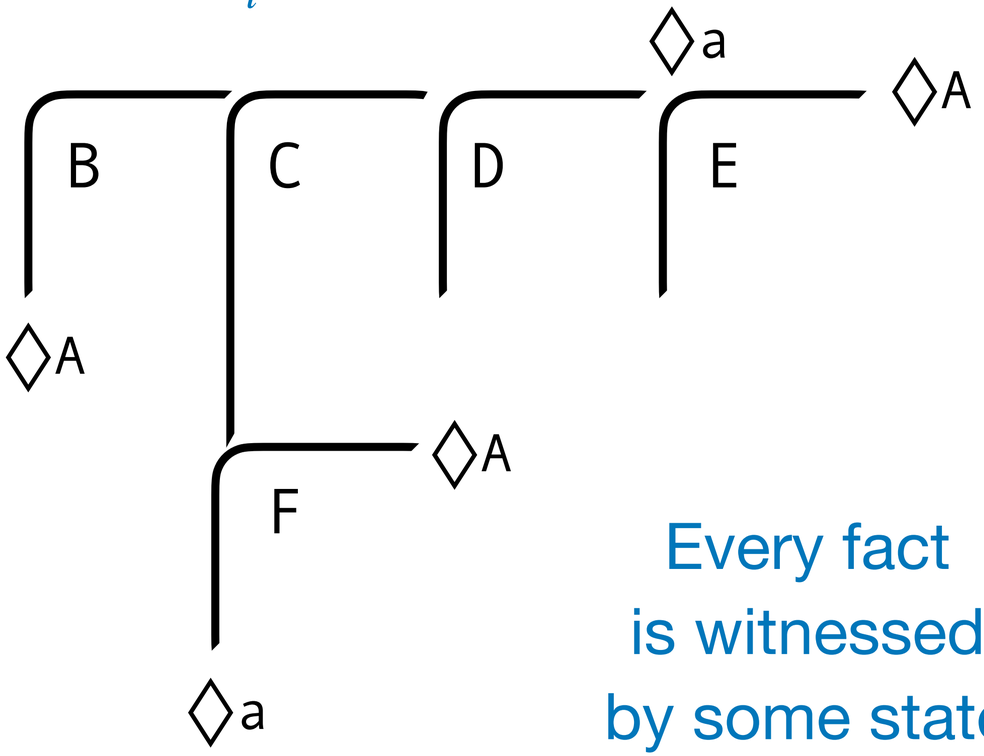
J' = HJ



Closed world assumption

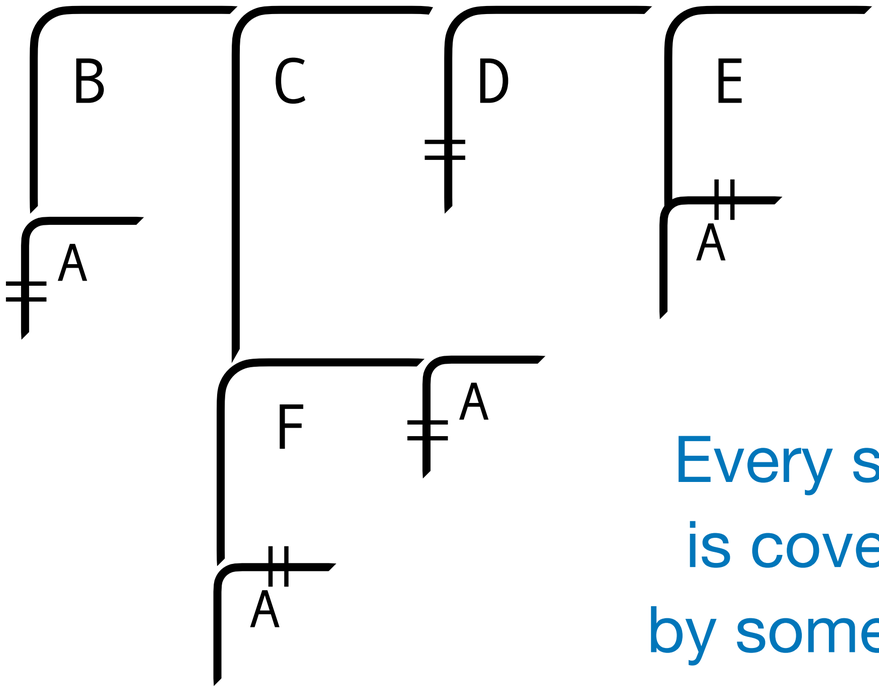
"absence of evidence is evidence of absence"

$$facts = \bigwedge_i \Diamond \alpha_i$$



Every fact
is witnessed
by some state

$$belief = \Box \bigcup_i \alpha_i$$



Every state
is covered
by some fact