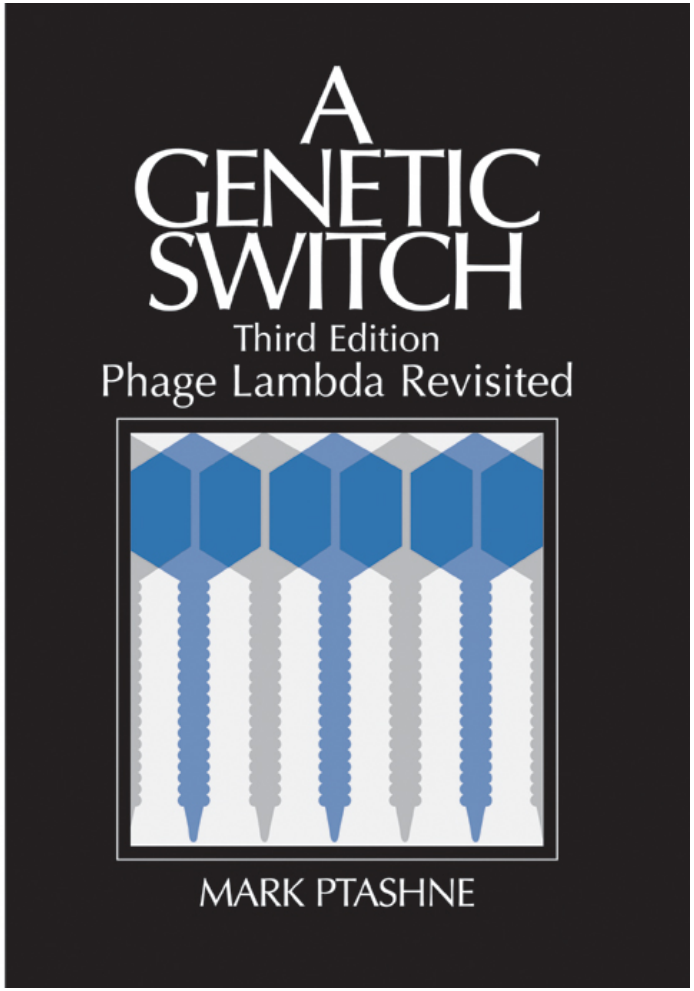
Natural processes and scientific reasoning
(sequentialization as the result of concurrent computation)

René Vestergaard

Pivots: 1) background, 2) focus, 3) example

1. a reasoning style is (often) a model of computation
2. a model of computation is mathematics $\rightsquigarrow$ can be wrangled
3. a particular computation has problem-specific import

“Raw data”: standard molecular-biology reference



“Get “A Genetic Switch, 3rd”. Read it and read it again.”

Review on amazon.com by “Ardent Reader”

Informally: molecular coding of physiology

TRICKY: getting it exactly right!

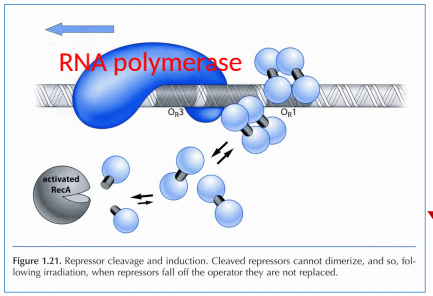


Figure 1.21. Repressor cleavage and induction. Cleaved repressors cannot dimerize, and so, following irradiation, when repressors fall off the operator they are not replaced.

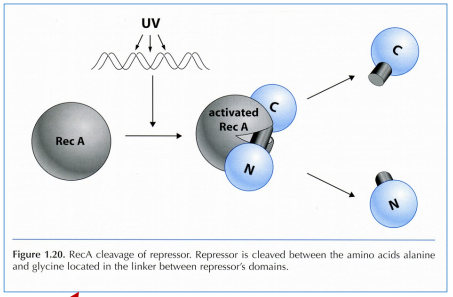
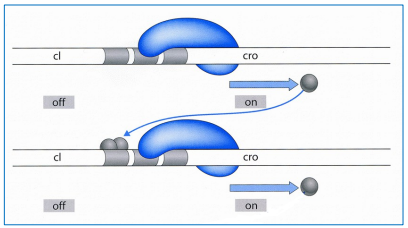
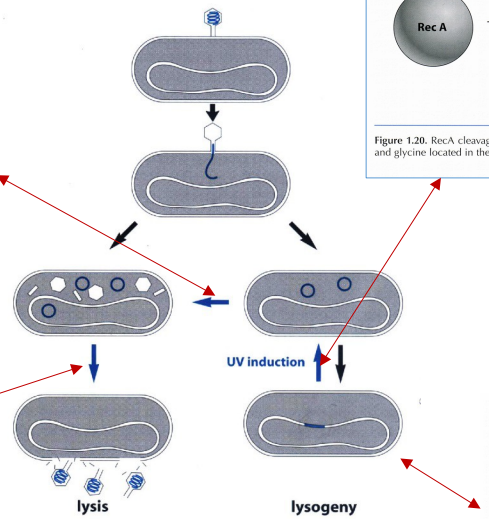


Figure 1.20. RecA cleavage of repressor. Repressor is cleaved between the amino acids alanine and glycine located in the linker between repressor's domains.





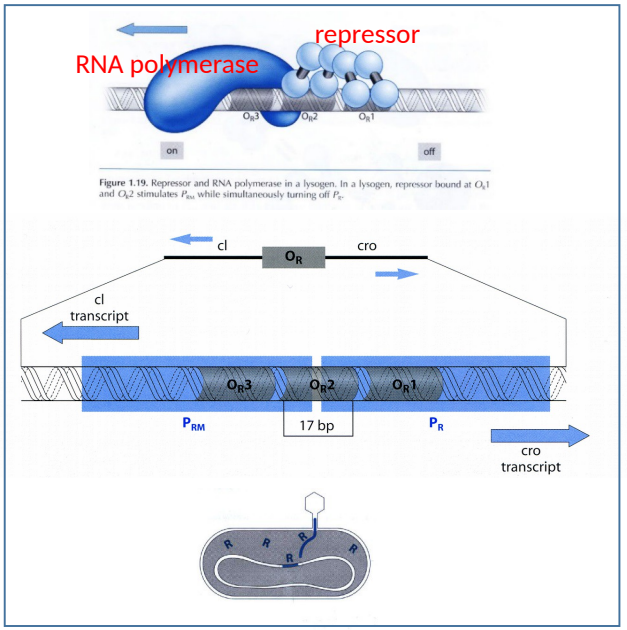
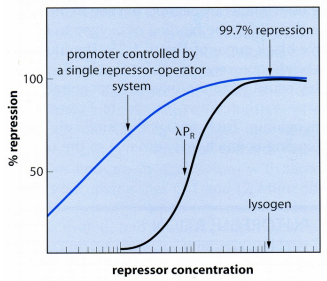
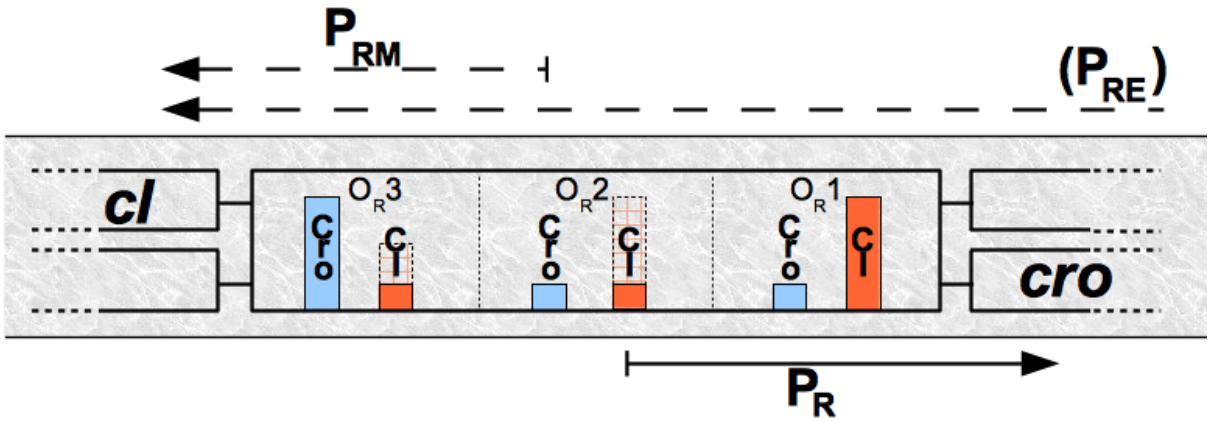
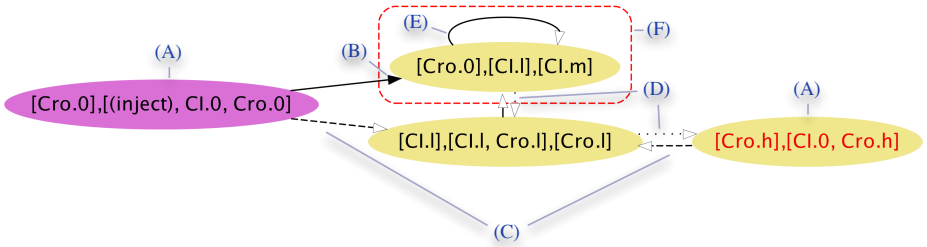


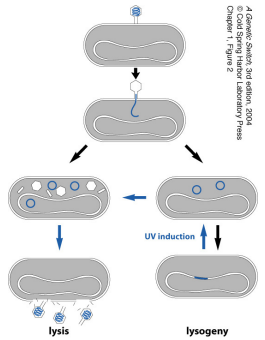
Figure 1.19. Repressor and RNA polymerase in a lysogen. In a lysogen, repressor bound at O_R1 and O_R2 stimulates P_{RM}, while simultaneously turning off P_R.

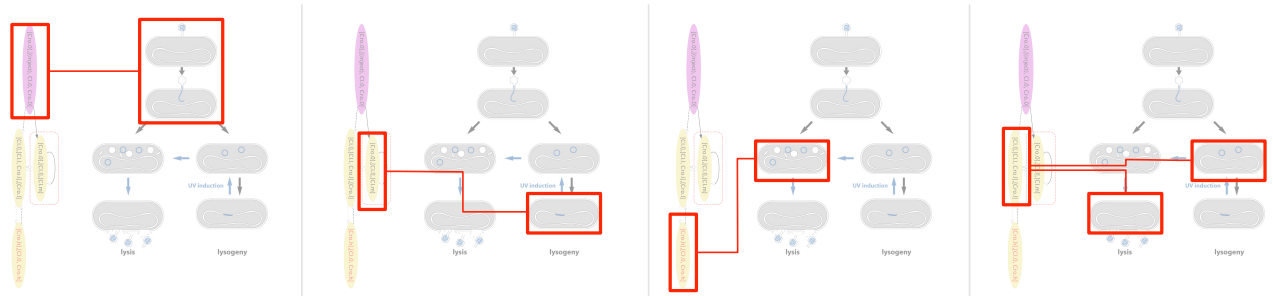


Formally: start and end point (but middle is crux)









Formally: intermediate coextension logic over EPPCs

Key step in augmented (inverted) hierarchy of reductionist reasoning style

Definition (*may* happen in compartment, under genotype g)

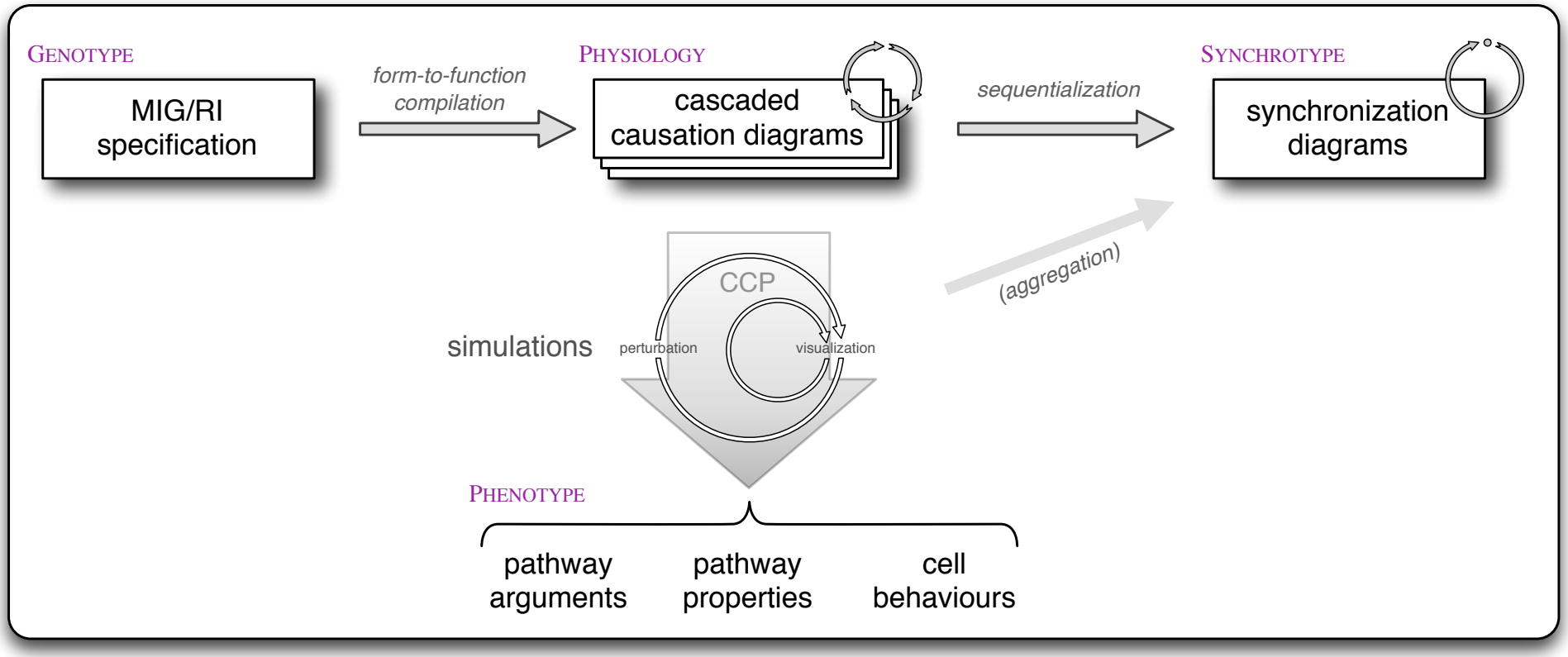
$$\frac{}{g \vdash A \overset{\varepsilon}{\rightsquigarrow} B}\ \text{(interference)} \qquad \text{if} \begin{cases} \exists \vec{R}_i, \vec{M}_i, \vec{P}_i, \vec{I}_i \quad . \\ \forall i \,.\, R_i \xrightarrow[I_i]{M_i} P_i \in \text{EPPC}(g) \\ \wedge\ \forall i \,.\, R_i \cup M_i \subseteq A \\ \wedge\ \forall i \,.\, I_i \cap A = \emptyset \\ \wedge\ \text{CoinhFree}(\{R_i \xrightarrow[I_i]{M_i} P_i\}_i) \\ \wedge\ B = (A \setminus \bigcup_i R_i) \cup \bigcup_i P_i \\ \wedge\ \text{well-formed}(B) \end{cases}$$

$$\frac{g \vdash A \overset{\tau_1}{\rightsquigarrow} B \quad g \vdash B \overset{\tau_2}{\rightsquigarrow} C}{g \vdash A \overset{\tau_1 B \tau_2}{\rightsquigarrow} C}\ \text{(sequence)}$$

CEqEA: Ptashne-style reasoning as computation

demonstration

+



Coextension reasoning example

Pathway argument [informal inference]

"Two changes result [from SOS/RecA]. First, as repressor vacates $O_R\{1,2\}$ the rate of repressor synthesis drops (because repressor is required to turn on transcription of its own gene); and second, polymerase binds to P_R to begin transcription of cro."* [24:-7]

Activity trace [formal inference, with all pertinent details]

```
1  [CI.1, Cro.0, RecA.*] + [DNA.ss]
2  ->  [SOS]
3  x>  ![] |-- [DNA.ss]  ;  [PR_tr] |-- [CI.1!@[OLR12, OLR123]]  ;
       [OLR12_CI, PRM_tr] |-- [SOS]

5  [Cro.0, RecA.*] + [CI.0, DNA.ss]
6  ->  [PR_tr]  ;  [SOS]
7  x>  ![] |-- [DNA.ss]

9  [Cro.1, RecA.*] + [CI.0, DNA.ss]
10 ->  [PR_tr]  ;  [SOS]
11 x>  ![] |-- [DNA.ss]
```

$\lambda^{[AGS3]}$ ab-intra phenotype: [AGS3] pathway properties

pathway argument < *pathway property* < cell behaviour

Key: ♡ stands for "absent a host SOS response" and ¬♡ stands for its negation "during a host SOS response". ♠ stands for "with CII below high concentration and with CI below physiological concentration" and ¬♠ stands for its negation "with CII at high concentration or with CI at physiological concentration"; *italics* not in [AGS3].

0/ *lysogeny's means, when viable (♡ and ¬♠), take precedence over a lytic attack's*

1/ CII determines λ's pathway at fresh infection ♡: lysogeny at high vs lytic attack below

2/ the lysogenic cycle is homeostatic, i.e., is sustainable and absorbing perturbations

- a/ protein concentrations are CI controlled in the lysogenic cycle
  - α/ an initial CI concentration is and must be established by high CII ♡
  - β/ once established, CI remains physiological and Cro non-physiological ♡
- b/ a switch to the lytic attack is not effected by
  - i/ natural or operative means ♡ — this is the basis of sustainability
  - ii/ Cro-perturbation ♡ — this is the key means of homeostasis
  - iii/ super-infection ♡ — this is λ-immunity
- c/ a switch to the lytic attack is effected by UV-irradiation (via ¬♡)

3/ the lytic attack expires the host and, else, is sustainable but *not homeostatic*

- a/ protein concentrations are programmatic during a lytic attack
  - α/ the lytic attack is constitutive ¬♡ or ♠
  - β/ once initiated, Cro remains physiological and CI inoperative ¬♡ or ♠
  - γ/ a lytic attack expires the host after Cro has reached high ¬♡ or ♠
- b/ *ignoring host expiration, a switch to the lysogenic cycle is not effected by*
  - i/ *natural or operative means ¬♡ or ♠* — this is the basis of sustainability
  - ii/ CI-perturbation ¬♡, although the lytic attack may be set back
  - iii/ *super-infection ¬♡ or ♠* — not quite anti-immunity, see Chp.S5
- c/ *ignoring host expiration, a switch to lysogeny may be effected by*
  - ii/ *CI-perturbation ♡ (and ♠ in order for the lytic attack to be viable, see 3/d/)*
- d/ a lytic attack is unlikely ♡ and ¬♠

Retrodiction, based also on non-observable properties

Retrodiction (I) Fig.6 0/ is not considered in [AGS3] and is, in fact, contradicting superseded statements from earlier editions and printings, e.g., *"Cro [being transcribed from P_R] determines the course of events"* [24:-5] and *"the first Cro to be synthesized binds to O_R3 [which] prevents [RNAP] from binding to P_{RM} and abolishes further synthesis of repressor. At this point the switch has been thrown and lytic growth ensues."* [25:2a,2b,3a,3b]. 'Pathway property' 0/ is consistent with all parts of [AGS3] that are not explicitly stated or implied to be superseded by the recent findings that prompted the 3rd edition of the monograph, see *"[t]his new edition is prompted by discoveries [...] that add to, rather than reformulate, the earlier story"* [xi:11] and *"the conclusion that Cro must bind O_R3 to trigger the transition to lytic growth, although not excluded, remains uncertain"* [121:-4]. In particular, see *"If repressor were added to a phage beginning its lytic cycle, growth would be inhibited"* [62:-1]. More, 1/, 2/a/α/, 2/b/ii/, 3/c/ii/, and 3/d/ would not hold in their present form if 0/ did not hold, and, as we show in Sect.S5.4/Cro, failure of 0/ results in anti-immunity. [...] Our $\lambda^{[AGS3]}$ is a concrete model of λ where Cro binding to O_R3 is not needed to trigger the switch, see [121:-4].

Retrodiction (II) A main implication of the structure of the 'pathway properties' we establish, see Fig.6, is that lysogenic cycles and lytic attacks (absent host expiration) exhibit fundamentally different stability properties, which appears to be a novel inference (of a probably-known observation, see [92:Fig.4.23:1]). [...]

Retrodiction (III) A further implication of the 'pathway properties' we establish, see Fig.6, is that the lysogenic-to-lytic switch is effected by activated-RecA-mediated CI-proteolysis at the onset of the switch and only later affected also by Cro inhibition of *cI* transcription, see 2/c/, 3/c/ii/, and the varying of Cro's intrinsic affinities in Chp.S5. [...]

Retrodiction (IV) Fig.6 and Chp.S3 do not firmly establish at what point a lysogenic-to-lytic switch can be said to have been initiated, see [121:-4]. The clear suggestion is that it is at the physiologically-zero concentration of CI at which constitutive *cro*-transcription becomes physiological, see 0/, 2/a/α/ vs 3/a/α/, and Retrodiction (III). A consequence of this, see 2/a/ vs 3/a/β/, is the known property that there is no turning back once the lytic attack is under way, at least in the absence of physiological CII and of stochastic CI that is sufficiently strong to re-establish RNAP-recruitment at P_{RM}, see 3/c/ii/ and 3/d/.

Meta-theory guarantee 1/2: the methods are logical

Definition

$$R \xrightarrow[I]{M} P/A \quad \triangleq \quad \begin{cases} A \overset{\varepsilon}{\rightsquigarrow} (A \setminus R) \cup P & \text{if } R \cup M \subseteq A \wedge I \cap A = \emptyset \wedge \ldots \\ A \overset{\varepsilon}{\rightsquigarrow} A & \text{o/w} \end{cases}$$

Proposition (Unit)

$$\forall g, p, A \quad . \quad p \in EPPC(g) \wedge \ldots \Rightarrow g \vdash p/A$$

Proposition (Recompose)

$$\begin{aligned} \forall g_1, g_2, p, A \quad . \quad & (\forall p_x, X \quad . \quad p_x \in EPPC(g_1) \wedge \ldots \Rightarrow g_2 \vdash p_x/X) \\ & \wedge\, g_1 \vdash p/A \wedge \ldots \\ & \Downarrow \\ & g_2 \vdash p/A \end{aligned}$$

Meta-theory guarantee 1/2: the methods are logical

Definition

$$R \xrightarrow[I]{M} P/A \quad \triangleq \quad \begin{cases} A \overset{\varepsilon}{\rightsquigarrow} (A \setminus R) \cup P & \text{if } R \cup M \subseteq A \wedge I \cap A = \emptyset \wedge \ldots \\ A \overset{\varepsilon}{\rightsquigarrow} A & \text{o/w} \end{cases}$$

Proposition (Unit)

$$\forall g, p, A \quad . \quad p \in EPPC(g) \wedge \ldots \Rightarrow g \vdash p/A$$

Proposition (Recompose)

$$\begin{aligned} \forall g_1, g_2, p, A \quad . \quad & (\forall p_x, X \quad . \quad p_x \in EPPC(g_1) \wedge \ldots \Rightarrow g_2 \vdash p_x/X) \\ & \wedge\, g_1 \vdash p/A \wedge \ldots \\ & \Downarrow \\ & g_2 \vdash p/A \end{aligned}$$

Theorem ((cut) is admissible)

$$\frac{g \vdash p_0/A \quad p_0{+}{+}g \vdash p/A}{g \vdash p/A}\ (cut) \qquad \text{if } \ldots$$

Meta-theory guarantee 2/2: the methods do something

Proposition (Non-empty)

$$\exists g, A, B, \tau \quad . \quad g \vdash A \overset{\tau}{\rightsquigarrow} B$$

Theorem (Non-trivial/consistent)

$$\exists A, B, \tau \quad . \quad \bot \nvdash A \overset{\tau}{\rightsquigarrow} B$$

Sequentialization, trace-monoid style

Definition (Synchrotype [idealized])

By *synchrotype*, we mean an account of how a collection of physiology-change processes may interleave with each other under coextension where we do not distinguish different execution orders of processes that are independent of each other and environmental factors are retained as conditionals.

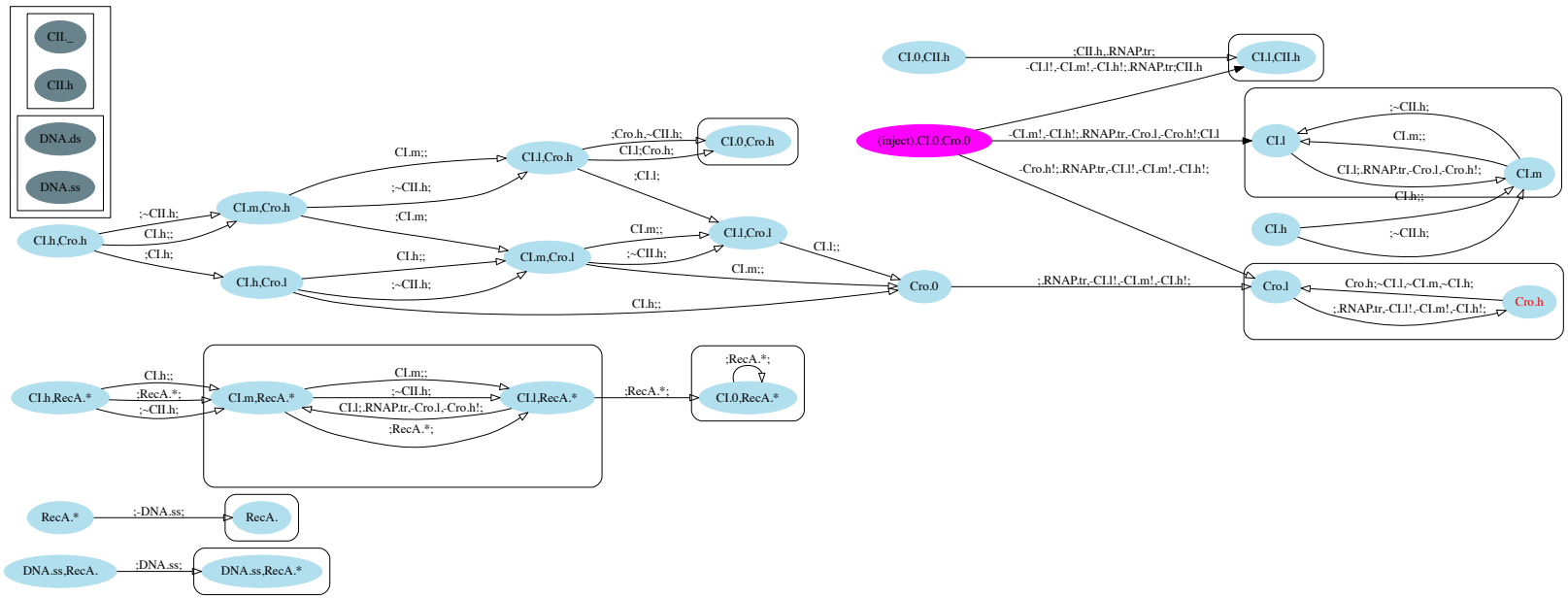
Technically

CCP traversals of cascaded causation diagram form a partially-commutative monoid, with composition. Synchrotype is (essentially) the trace monoid = independence factored out: leaves graph where edges are changes breaking synchronicity.

Independence 1/2: interpolation of observable values

Axiom [informal principle]

"As [Cro] binds it turns off [P_R], but as the cells grow and divide, the concentration of [Cro] drops, and [P_R] turns on again. A steady state is reached at which the rate of synthesis of [Cro] just balances its rate of dilution and, presumably, a constant concentration of [Cro] is maintained. In this situation [Cro] diminishes, but does not abolish, its own synthesis." [92:-3]



Independence 1/2 (cont'd): sustaining

Definition (A/T/N-sustainers)

States may be Always present/Transitory/Never present.

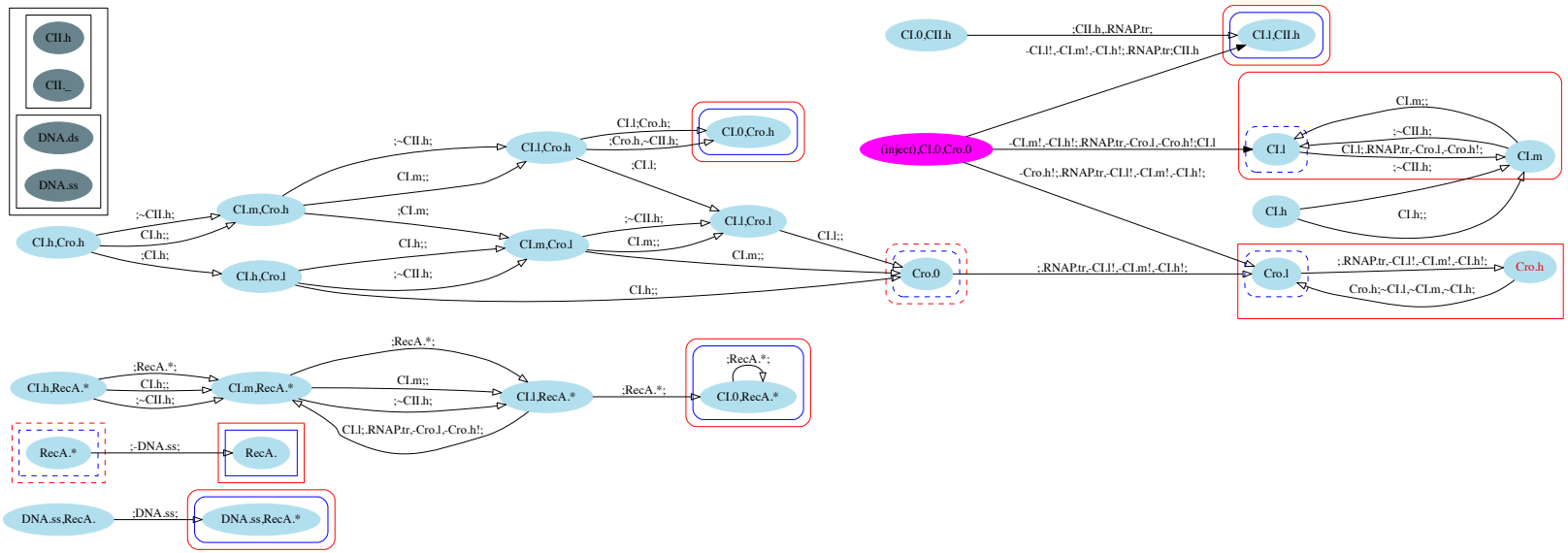
Definition (A/T/N-sustained equilibria)

Given a classification of the considered states as A/T/N sustainers, we first find all strongly-connected components over the graph without A-inhibited edges. We then classify any found component as a *sustained equilibrium* of the given A/T/N type if it has no out-edges in the graph without A- and/or T-inhibited edges. By default, we consider only direct inhibitors for sustaining: nested inhibition tends to express precedence.

Definition (Min/Max-sustained equilibria)

- Sustained equilibria with all states classified as transitory are referred to as *max equilibria* and are annotated to cascaded causation diagrams as red boxes.
- Sustained equilibria with all states classified as always present are referred to as *min equilibria* and are annotated to cascaded causation diagrams as blue boxes.

Id dependence 1/2 (cont'd): min/max sustaining



CEqEA cascaded causation (v28/e39/c8; i_/x9/a_/n9)
lambda[AGS3]_RI
ctxt: RNAP.tr

Independence 2/2: sustained-equilibria coexistability

Definition (Reconcilable)

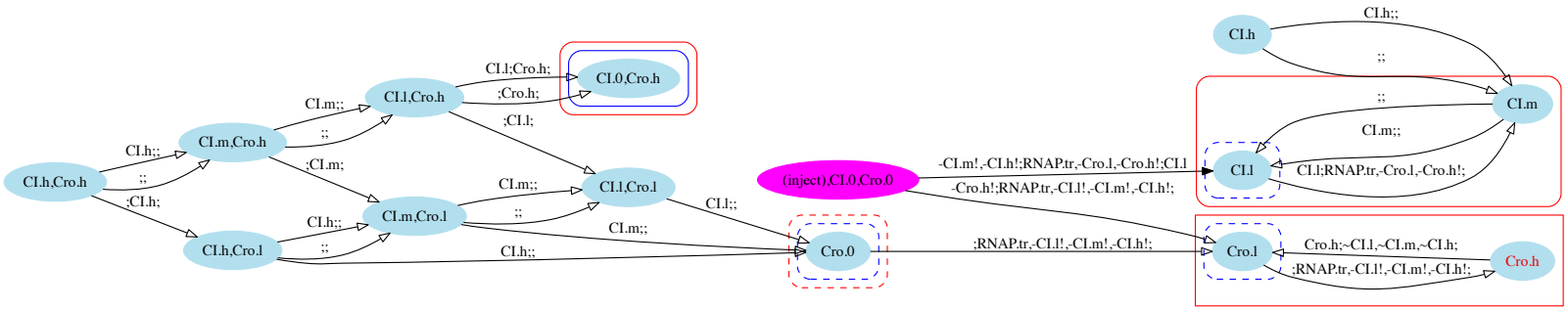
Each node in a sustained equilibrium must be able to pair up with a node from each of the other sustained equilibria without having some schema instance be in two states.

⋮

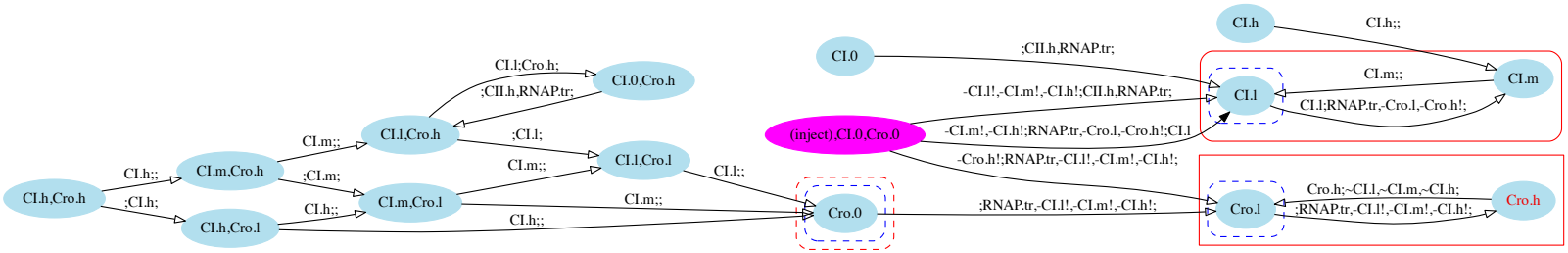
Definition (Orthogonal — USED HERE!)

A collection of sustained equilibria are orthogonal if no schema instance is in different states in (strictly) different equilibria and if no sustained equilibrium loses strong-connectivity under inhibition by the content of the other sustained equilibria.

Targeted cascaded causation diagrams: exogeny

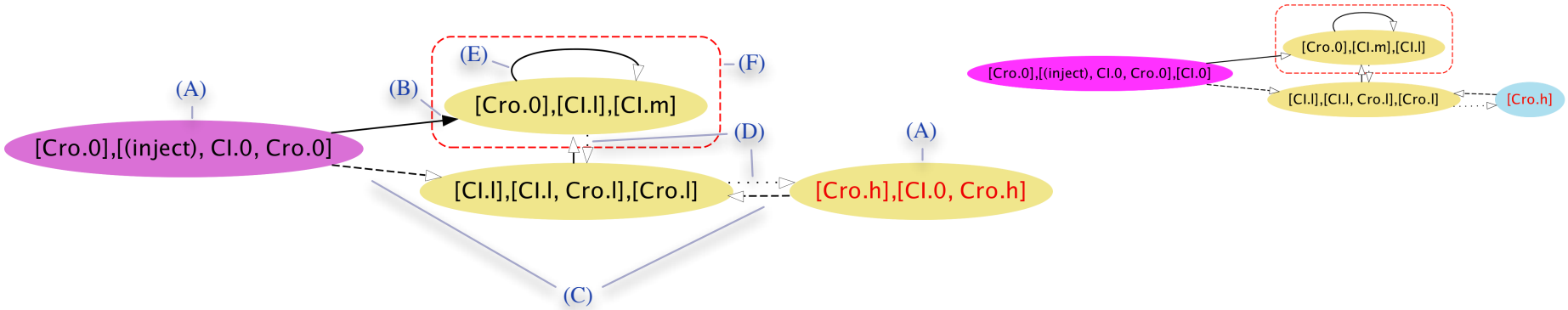


CEqEA cascaded causation (v14/e26/c1; i_/x4/a_/n4)
lambda[AGS3]_RI
ctxt#1: [CII._]; RNAP.tr



CEqEA cascaded causation (v15/e22/c1; i_/x3/a_/n3)
lambda[AGS3]_RI
ctxt#2: [CII.h]; RNAP.tr

$\lambda^{[AGS3]}$ synchronization diagrams, w/molecular details



(A) Synchronization points are groups of indistinguishable causation nodes, with expiration (red text) and channels (magenta background) lifted from the causation level. The "synchro-expiration" is terminal only in that sense, meaning lysis is not regulatorily terminal, see Retrodiction (II). The "synchro-channel" contains both inter- and intra-nodes, meaning it combines the two top nodes in Ptashne's diagram.

(B) Edges with a filled arrow head are subject to the TargetRequired condition: some requisite mediator is present only in the target node. This implies that the direct route to λ-lysogeny is available only to a super-infection in the considered case, namely with CII below high concentration, see 1/ and 2/b/iii/. The corresponding edge in the synchronization diagram with CII at high concentration is not TargetRequired, see Sect.S4.4, meaning fresh λ-lysogeny is possible in that case, see 2/a/α/.

(C) Dashed edges are subject to inhibition in the target node. CI is an inhibitor for the left edge, meaning only continued λ-lysogeny is possible in case of super-infection, see 2/b/iii/. For the right edge, see Retrodiction (III).

(D) Dotted edges are subject to inhibition in the source node. The inhibitor here is CI, meaning a lytic attack requires CI to become non-physiological by means other than what the λ-genotype codes for or by a (severe) stochastic event, see 2/{a,b}/ vs 2/c/.

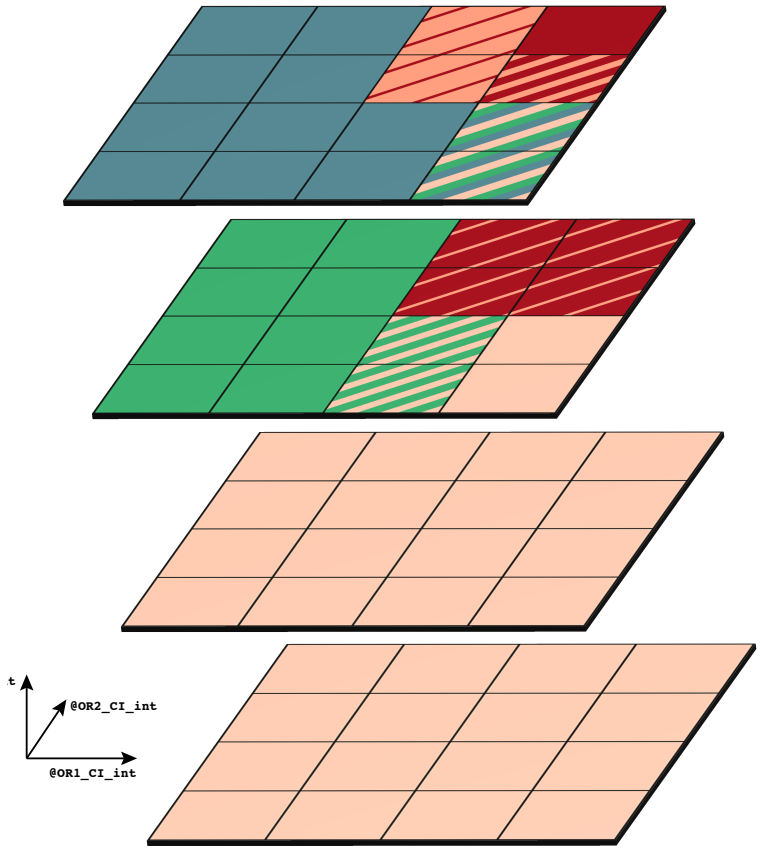
(E) Reflexive edges indicate self-regulation. Here, of λ-lysogeny, see 2/a/.

(F) Boxes indicate "synchro-sustainability", see Chp.S4. Here, of λ-lysogeny, while all other synchronization points are likely to be transient.

(D)–(F) jointly predict that the λ-lysogeny box will be relatively stable (in vacuo). Combined with the specifics of the out-edges of the intermediate synchronization point, we can further predict that λ-lysogeny will be homeostatic (in situ): if we are pushed out of the box, the only option will be to go back into the box in most circumstances.

Combinatorial: pheno-over-genotype phase space

phenotype classification	$[\_]^{\succ}$	$[\mathrm{h}]^{\succ}$	$[\_]^{\curlyvee}_{1_0}$	$[\_]^{\curlyvee}_{2_0}$	$4^{\succ}_0$	$(\cdot)^{\succ}_{+}$	$(\cdot)^{\succ}_{\mathrm{UV}}$
λ*virulent*	$\otimes$	$\otimes$	$\ldots\otimes$	$\ldots\otimes$	$\otimes$		
λ*virulent/clear*$^+$	$\otimes$	$\otimes$	$\ldots\otimes$	$\ldots 0$	$\otimes$		
λ*clear*$^+$*;virulent*	$\otimes$	$\otimes$	$\ldots\otimes$	$\ldots 0$	$\{3,2\}_0;\otimes$	$\shortparallel;\otimes$	$\otimes$
λ*clear*$^+$*/virulent*	$\otimes$	$\otimes$	$\ldots\otimes$	$\ldots 0$	$\{3,2\}_0$	$\shortparallel$	$\otimes$
λ*clear*$^+$	$\otimes$	$\otimes$	$\ldots 0$	$\ldots 0$			
λ*clear*$^-$	$\otimes$	$0_0,1_1$	$\ldots 0$	$\ldots 0$	$\otimes$		
λ*temperate;clear*$^-$	$\otimes$	$0_0,1_1$	$\ldots 0$	$\ldots 0$	$\{2,1\}_0;\otimes$	$\shortparallel;\otimes$	$\otimes$
λ*super;temp;clear*$^-$	$\otimes$	$0_0,1_1$	$\ldots 0$	$\ldots 0$	$\{3,2,1\}_0;\otimes$	$\shortparallel;\otimes$	$\otimes$
λ*temperate*	$\otimes$	$\{2,1\}_0$	$\ldots 0$	$\ldots 0$	$\{2,1\}_0$	$\shortparallel$	$\otimes$
λ*super-temperate*	$\otimes$	$\{3,2\}_0$	$\ldots 0$	$\ldots 0$	$\{3,2\}_0$	$\shortparallel$	$\otimes$



Compound phenotypes

With '/' — tend to left, may be right

With ':' — non-deterministic choice

Main topics

[*Principia Mathematica*]: "Principles of molecular-biology reasoning: *A Genetic Switch* is formally correct"

[*Curry-Howard*]: "Mechanisms of molecular-biology reasoning: *A Genetic Switch* 'fits together' *phage* λ"

[Proof assistant]: "Formally-verified cause-and-effect reasoning for emergent properties"

[Deep structure]: "Combinatorial basis of morphogenesis: elementary processes of digitation"

[Computation]: "Synchrotyping"