Dobbiaco Lectures 2010 (1)
Solved and Unsolved Problems in Biology

Bud Mishra

Room 1002, 715 Broadway, Courant Institute, NYU, New York, USA

Dobbiaco

Outline

PART I : Narratives

Main theses

"...

1. ...
2. "**Where (or of what) one cannot speak, one must pass over in silence.**"

–Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, 1921.

Outline

Information Processing inside a Cell

- Biology — A study of certain special kinds of information processing systems. Or, multi-agent repeated game working under some *replicator dynamics*.
- This view disregards most of what biologists (still) study: *biochemistry*, *molecular biology*, *cell biology*, etc. as these can *only* lead up to an understanding of the structural machinery underlying the biological systems.

Genomes

- Hereditary information of an organism is encoded in its DNA and enclosed in a cell (unless it is a virus).
- All the information contained in the DNA of a single organism is its *genome*.
- Understanding information encoding in DNA: Envision a DNA molecule to be just a *very* long sequence of *nucleotides* or *bases*:

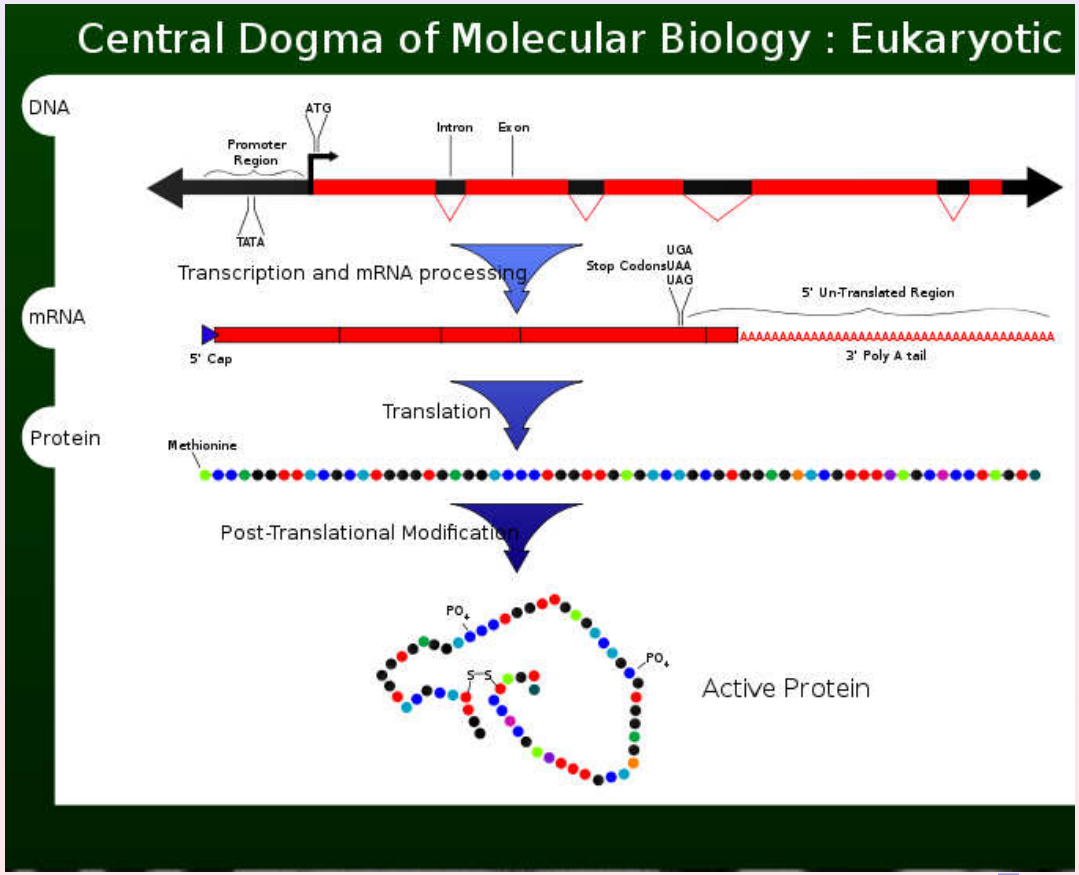
$$\Sigma = \{\mathbf{A}, \mathbf{T}, \mathbf{C}, \mathbf{G}\}.$$

Central Dogma

- The intermediate molecule carrying the information out of the nucleus of an eukaryotic cell is RNA, a single stranded polymer with the same bases as DNA except the base *thymine* is replaced by *uracil*, **U**.
- RNA also controls the translation process in which amino acids are created making up the proteins.
- The central dogma (due to Francis Crick in 1958) states that these information flows are all unidirectional...

Central Dogma

" The central dogma states that once 'information' has passed into protein it cannot get out again. The transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein, may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information means here the precise determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein ."


Central Dogma of Molecular Biology : Eukaryotic
DNA
ATG
Promoter Region
Intron
Exon
TATA
Transcription and mRNA processing
UGA
Stop CodonsUAA
UAG
5' Un-Translated Region
mRNA
5' Cap
3' Poly A tail
Translation
Protein
Methionine
Post-Translational Modification
PO4
S—S
Active Protein

Genes

- A specific region of DNA that ultimately determines the synthesis of proteins (through the transcription and translation) is called a *gene*.
- Transcription of a gene to a *messenger RNA* is keyed by an RNA polymerase enzyme, which attaches to a *core promoter* (a specific sequence adjacent to the relevant structural gene).

Gene Regulation

- Regulatory sequences such as *silencers* and *enhancers* are responsible in controlling the rate of transcription by their influence on the RNA polymerase.
- Regulation involves a feedback control loop involving many large families of *activator* and *repressor* proteins that bind with DNA
- These in turn, transpond the RNA polymerase by *coactivator proteins* and *basal factors*.

Gene Regulation

- **Question 1**: *How do we know anything about the architecture of transcriptional regulation??*
- Especially, since genes are regulated in a manner that is rather dispersed and fairly complicated. *It is believed that...*
  - The enhancer and silencer sequences occur over a wide region spanning many Kb's from the core promoter on either directions.
  - A gene may have many silencers and enhancers and can be shared among the genes.
  - They are not unique—different genes may have different combinations

Gene Transcription

- The transcription of DNA in to m-RNA is performed with a single strand of DNA (the sense strand) around the region corresponding to a gene.
- The double helix untwists momentarily to create a transcriptional bubble which moves along the DNA in the 3' - 5' direction (of the sense strand) as the complementary m-RNA synthesis progresses adding one RNA nucleotide at a time at the 3' end of the RNA, attaching a **U** (respectively, **A**, **G** and **C**) for the corresponding DNA base of **A** (respectively, **T**, **C** and **G**).

Gene Transcription

- The transcription process ends when a special sequence called the *termination signal* is encountered.
- This newly synthesized m-RNA are capped by attaching special nucleotide sequences to the 5' and 3' ends. This molecule is called a *pre-m-RNA*.
- In eukaryotic cells, the region of DNA that is transcribed into a pre-m-RNA involves more than just the information needed to synthesize the proteins.

Gene Transcription

- The DNA subsequences that contain the information or *code* for protein (somewhat indirectly) are the so-called *exons* which are interrupted by regions of *introns*, the non-coding regions.
- Note that pre-m-RNA contains both exons and introns and needs to be altered to excise all the intronic subsequences in preparation for the translation process—this is done by the *spliceosome*.

Translation of a Gene

- The translation process begins at a particular location of the m-RNA called the translation start sequence (usuall **AUG**) and is mediated by the *transfer RNA* (t-RNA), made up of a group of small RNA molecules, each with specifity for a particular amino acid.
- The t-RNA's carry the amino acids to the *ribosomes*, the site of protein synthesis, where they are attached to a growing polypeptide. The translation stops when one of the three trinucleotides **UAA**, **UAG**, **UGA** is encountered.

Codons

- Each 3 consecutive (nonoverlapping) bases of m-RNA (corresponding to a *codon*) codes for a specific amino acid. There are $4^3 = 64$ possible trinucleotide *codons* belonging to the set

$$\{\mathbf{U}, \mathbf{A}, \mathbf{G}, \mathbf{C}\} \times \{\mathbf{U}, \mathbf{A}, \mathbf{G}, \mathbf{C}\} \times \{\mathbf{U}, \mathbf{A}, \mathbf{G}, \mathbf{C}\}.$$

- The codon **AUG** is the *start codon* and the codons **UAA**, **UAG**, **UGA** are the *stop codons*.

Codons

- The line of nucleotides between and including the start and stop codons is called an *open reading frame* (ORF) and one can assume that all the information of interest to us resides in the ORF's.
- The mapping from the codons to amino acid (and naturally extended to a mapping from ORF's polypeptides by a homomorphism) given by

$$\begin{aligned} F_P : \{\mathbf{U}, \mathbf{A}, \mathbf{G}, \mathbf{C}\}^{\mathbf{3}} &\rightarrow \{A, R, D, N, C, E, Q, G, H, \\ &\qquad I, L, K, M, F, P, S, T, W, Y, V\} \\ \mathbf{UUU} &\mapsto F(= \text{Phe} = \text{phenylamine}) \end{aligned}$$

Coding of the Amino Acids

	RF_1				
RF_0	**G**	**A**	**C**	**U**	RF_1
G	Gly	Glu	Ala	Val	**G**
	Gly	GLu	Ala	Val	**A**
	Gly	Asp	Ala	Val	**C**
	Gly	Asp	Ala	Val	**U**
A	Arg	Lys	Thr	Met	**G**
	Arg	Lys	Thr	Ile	**A**
	Ser	Asn	Thr	Ile	**C**
	Ser	Asn	Thr	Ile	**U**
C	Arg	Gln	Pro	Leu	**G**
	Arg	Gln	Pro	Leu	**A**
	Arg	His	Pro	Leu	**C**
	Arg	His	Pro	Leu	**U**
U	Trp	Stop	Ser	Leu	**G**
	Stop	Stop	Ser	Leu	**A**
	Cys	Tyr	Ser	Phe	**C**
	Cys	Tyr	Ser	Phe	**U**

Genetic Codes

- The genetic code for each triplet can be read of by looking at the entry given by the first letter (RF_0, base in the reading frame 0) along the left column, the second letter (RF_1, base in the reading frame 1) along the row and the third letter (RF_2, base in the reading frame 2).
- In an ORF, a given occurrence of a base is said to be in *reading frame* 0, 1, or 2, if it is the first, second or third letter in a codon, respectively.
- A codon is said to be *in-frame* if its first base is in reading frame 0.

Cell as an Automaton

- The set of mRNA's in the cell (different species and their copy numbers) determines a cell's **state**.
- The translated proteins and their copy numbers are determined by the various mRNA's (i.e., cell's states).
- Some of the proteins are **transcriptional factors** and determine which (and how fast) various mRNA's are to be transcribed; thus, they determine the **transitions to the next state**.
- The state transition may not be deterministic; the randomness involved in probabilistic transitions is a property of various cellular conditions, protein copy number, epigenetics, etc.

Cell's States are Hidden

- We have no viable technology to observe cell's states (e.g., all the mRNA's, miRNA's, proteins [with their post-translational modifications], etc.) ...
- We can observe some of its outputs: Some of the known genes, surface proteins or proteins that can be tagged with fluorescent probes...
- The observed outputs are given by a function of the (hidden) states; this function is also subject to some randomness (due to the nature of the underlying biology)...
- Often, we will want to understand the dynamics of the system from the observed outputs...

PART II : Reasoning

Outline

Necessary and Sufficient Conditions

- **Brian Skyrms**:
- Practical application of knowledge of causes consists either in (i) producing the cause in order produce the effect or in (ii) removing the cause in order to prevent the effect.
- The word “cause” is used to mean several different things... It is more useful to talk about *necessary conditions* and *sufficient conditions*

Necessary and Sufficient Conditions

Definition

A property F is a *sufficient condition* for a property G if and only if *whenever F is present, G is present.*

Definition

A property H is a *necessary condition* for a property I if and only if *whenever I is present, H is present.*

- Whenever we say that A causes B, we sometimes mean that A is a sufficient condition for B, sometimes that A is a necessary condition for B, sometimes that A is both necessary and sufficient for B, and sometimes none of these things.

Necessary and Sufficient Conditions

Theorem

1. *If property A is a* sufficient condition *for a property B, then B is a* necessary condition *for a property A.*

Theorem

2. *If property C is a* necessary condition *for a property D, then D is a* sufficient condition *for a property C.*

Necessary and Sufficient Conditions

Theorem

1'. If property A is a sufficient condition *for a property B, then* $\neg B$ *is a* sufficient condition *for a property* $\neg A$.

Theorem

2'. If property C is a necessary condition *for a property D, then* $\neg D$ *is a* necessary condition *for a property* $\neg C$.

Necessary and Sufficient Conditions

Theorem

1". If property A is a sufficient condition *for a property B, then ¬A is a* necessary condition *for a property ¬B.*

Theorem

2". If property C is a necessary condition *for a property D, then ¬C is a* sufficient condition *for a property ¬D.*

The Method of Agreement

- A, B, C ad D are *possible conditioning properties*.
- E is a *conditioned property*.

	A	B	C	D	E
occurrence (1)	P	P	P	A	P
occurrence (2)	P	A	P	P	P
occurrence (3)	A	P	P	A	P

- C is a necessary condition for E.
- Occurrence (1) shows that D cannot be a *necessary condition* for E. It can be eliminated.
- In the same manner, occurrences (2) and (3) eliminate B and A, respectively, as possible *necessary conditions*

The Inverse Method of Agreement

- A, B, C ad D are *possible conditioning properties*.
- E is a *conditioned property*.

	A	B	C	D	E
occurrence (1)	P	A	A	A	A
occurrence (2)	A	P	A	A	A
occurrence (3)	P	A	P	A	A

- D is a sufficient condition for E.
- Occurrence (1) shows that A cannot be a *sufficient condition* for E. It can be eliminated.
- In the same manner, occurrences (2) and (3) eliminate B and C, respectively, as possible *sufficient conditions*.

LAC OPERON

- The modern study of gene regulation was initiated in the 1950's by Francois Jacob, Jacques Monod, Andre Lwoff, and many workers at the Institute Pasteur in Paris.
- They recognized that bacteria, like higher organisms, regulate expressions of their genes.
- For instance, the colon bacteria, *E. coli* does not express all its genes all of the time... even though it is a single-cell and has no developmental process to undergo.
- Jacob and his colleagues studied the ability of *E. coli* to grow on a wide array of different sugars — glucose and lactose.

β-Galactosidase

- The gene encoding the enzyme β-galactosidase is silent until its substrate – lactose – is added to the medium.
- The gene is then turned on and the enzyme is synthesized.
- Initially, there was some confusion whether this was an example of *adaptation* (a population property) or *mutation* (an individual property).
- The process seems too fast to be either a mutation or adaptation...
- Were it a mutation, then the evolution was too *Lamarckian* to be true.

A Simple Model

- One could imagine a model in which *E. coli* alternated between two states $\langle \{S_1, S_2\}, S_1, \mathcal{T} \rangle$, the transition from S_1 to S_2 being controlled by the "guard condition:"

$$[\text{Glucose}] > [\text{Lactose}]$$

and the reverse transition by the "guard condition:"

$$[\text{Lactose}] > [\text{Glucose}] \ .$$

- However, the system acts somewhat differently (a process called, "diauxie") for this state-machine model to be true.

MIll-Skyrms Approach applied to Lac operon

- Diauxie (Meaning double growth)[1] is the phenomenon in which the *E. coli* consumes the glucose as its main carbon source, and switches to lactose (if present) when there is no glucose left.

	G	$\neg G \wedge L$	$Cons(G)$	$Cons(L)$
wildtype (1)	P	A	P	A
wildtype (2)	A	P	A	P

- $\neg G \wedge L \Rightarrow_N Cons(L)$
- Lac^-: This is a mutant which is unable to grow on Lactose.

[1] AUXESIA is the Greek goddess who grants growth and prosperity to the fields...

- $Lac^- \mapsto_{mut} Lac^+$
- Lac^+ consumes lactose.
- Lac^+ controls an enzyme β-galactosidase. It can be assayed with ONPG (o-nitrophenyl-β-D-galactosidase)

$$ONPG\ (colorless) + \beta - \text{galactosidase} \Rightarrow galactose + o - nitrophenol\ (yellow)$$

($LacZ$ gene makes the enzyme β-galactosidase.)

	$LacZ$	$Cons(L)$
Lac^+	< 0.1	A
Lac^+	100	P
Lac^-	100	A

- (i) $LacZ \Rightarrow_N Cons(L)$ and $LacZ \not\Rightarrow_S Cons(L)$...
- Not yet shown $\neg G \wedge L \Rightarrow_N LacZ$

- Other enzymes involved: *LacY* (β-galactoside permease) and *LacA* (β-galactoside transacetylase)
- Lac^- is a permease-mutant.

	LacY	*Cons*(*L*)
Lac^+	95	P
Lac^-	< 0.1	A

- (ii) $LacY \Rightarrow_N Cons(L)$ and $LacY \not\Rightarrow_S Cons(L)$...
- Similarly, (iii) $LacA \Rightarrow_N Cons(L)$ and $LacA \not\Rightarrow_S Cons(L)$...

LacZ + LacY + LacA

$[LacZ \wedge LacY \wedge LacA \Rightarrow_N Cons(L)]$ and $[LacZ \perp LacY \perp LacA]$

- One can build z^+y^-, z^-y^+ and z^+y^+ mutants to show that $LacZ \perp LacY$

	$LacZ$	$LacY$	$Cons(L)$
z^+y^+	100	95	P
z^+y^-	100	< 0.1	A
z^-y^+	< 0.1	95	A

$LacZ \not\Rightarrow_N LacY$ and $LacZ \not\Rightarrow_S LacY$;

$LacY \not\Rightarrow_N LacZ$ and $LacY \not\Rightarrow_S LacZ$.

More Mutants

- There are two more systems to study: I^+ (wildtype) and I^-
  – In I^-, $LacZ$ is expressed constitutively.

	G	$\neg G \wedge L$	$LacZ$
$I+$	P	A	A
$I+$	A	P	P
$I-$	P	A	P
$I-$	A	P	P

$$(\neg G \wedge L) \perp LacZ | \{LacI, O\}$$

	G	$\neg G \wedge L$	$LacZ$
$I+O^-$	P	A	P
$I+O^-$	A	P	P
$I-O^+$	P	A	P
$I-O^+$	A	P	P

- In I^+O^- and I^-O^+, $LacZ$ is expressed constitutively.

Diploid Mutant

- To analyze regulatory mutants of the lac operon, Jacob developed a system by which a second copy of the lac genes (*LacI* with its promoter, and *LacZYA* with promoter and operator) could be introduced into a single cell.
- A culture of such bacteria, which are diploid for the lac genes but otherwise normal, is then tested for the regulatory phenotype.
- In particular, it is determined whether LacZ and LacY are made even in the absence of any interference (e.g. treatment with IPTG).
- This experiment, in which genes or gene clusters are tested pairwise, is called a **complementation test**.

Diploid Mutant

	G	$\neg G \wedge L$	$LacI$	O_B	$LacZ$
$I^-O^+Z^+$	P	A	A	P	P
$I^-O^+Z^+$	A	P	A	P	P
$I^+O^-Z^+$	P	A	P	A	P
$I^+O^-Z^+$	A	P	P	A	P
$I^+O^+Z^-$	P	A	P	P	A
$I^+O^+Z^-$	A	P	P	P	A
$I+O^+Z^+; I^-O^+Z^+$	P	A	{P,A}	P	A
$I+O^+Z^+; I^-O^+Z^+$	A	P	{P,A}	P	P
$I+O^+Z^+; I^+O^-Z^+$	P	A	P	{P,A}	P
$I+O^+Z^+; I^+O^-Z^+$	A	P	P	{P,A}	P
$I+O^+Z^-; I^+O^-Z^+$	P	A	P	{P,A}	P
$I+O^+Z^-; I^+O^-Z^+$	A	P	P	{P,A}	P

- How do we interpret this?

Interpretation

$$\begin{aligned} \neg G \wedge L &\Rightarrow_N \neg LacI(trans) \\ LacI &\Rightarrow_S O_B(cis) \\ \neg O_B &\Rightarrow_N \begin{cases} LacZ \\ LacY \\ LacA \end{cases} \end{aligned}$$

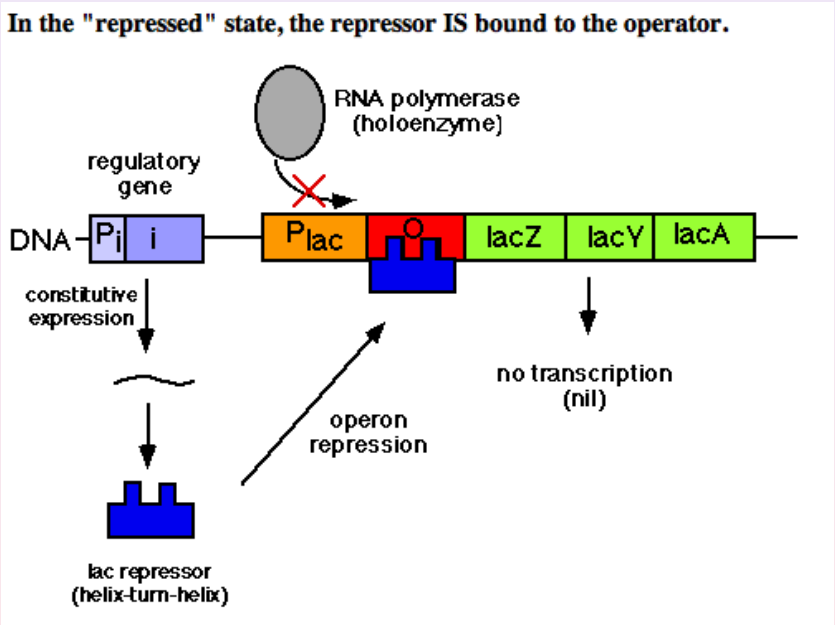
Interpretation

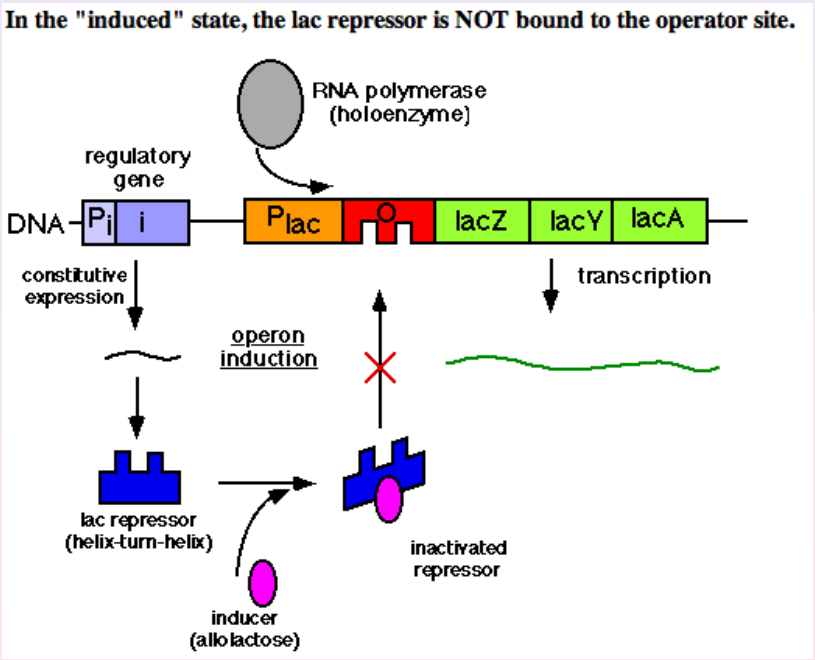
$$
\begin{aligned}
G \vee \neg L &\Rightarrow_S LacI \\
&\Rightarrow_S O_B \\
&\Rightarrow_S \begin{cases} \neg LacZ \\ \neg LacY \\ \neg LacA \end{cases}
\end{aligned}
$$

$$LacZ \wedge LacY \wedge LacA \Rightarrow_N Cons(L)$$

How the Story is Told

- The*lac operon* is an operon required for the transport and metabolism of lactose in *Escherichia coli* and some other enteric bacteria.
- It consists of three adjacent structural genes, a promoter, a terminator, and an operator. The lac operon is regulated by several factors including the availability of glucose and of lactose.
- Gene regulation of the *lac operon* was the first complex genetic regulatory mechanism to be elucidated and is one of the foremost examples of prokaryotic gene regulation.


In the "repressed" state, the repressor IS bound to the operator.
RNA polymerase
(holoenzyme)
regulatory
gene
DNA
Pi
i
Plac
O
lacZ
lacY
lacA
constitutive
expression
no transcription
(nil)
operon
repression
lac repressor
(helix-turn-helix)


In the "induced" state, the lac repressor is NOT bound to the operator site.
RNA polymerase
(holoenzyme)
regulatory
gene
DNA
Pi
i
Plac
O
lacZ
lacY
lacA
constitutive
expression
transcription
operon
induction
lac repressor
(helix-turn-helix)
inactivated
repressor
inducer
(allolactose)

How the Story is Told

- In its natural environment, *lac operon* is a complex mechanism to digest lactose efficiently.
- The cell can use lactose as an energy source, but it must produce the enzyme β-galactosidase to digest it into glucose. It would be inefficient to produce enzymes when there is no lactose available, or if there is a more readily-available energy source available (e.g. glucose).
- The lac operon uses a two-part control mechanism to ensure that the cell expends energy producing β-galactosidase, galactoside permease and thiogalactoside transacetylase only when necessary.

How the Story is Told

- The lac operon achieves this control with the **lac repressor**, which halts production in the absence of lactose, and the Catabolite activator protein (CAP), which assists in production in the absence of glucose.
- This dual control mechanism causes the sequential utilization of glucose and lactose in two distinct growth phases, known as **diauxie**.
- Similar diauxic growth patterns have been observed in bacterial growth on mixtures of other sugars as well, such as glucose and xylose or glucose and arabinose, etc. The genetic control mechanisms underlying such diauxic growth patterns are known as *xyl operon* and *ara operon*, etc.

Why was it so Hard to Interpret?

- For Hinshelwood the repressor (LacI) did not exist It was a misinterpretation:
- "*If we accept the view that all other effects on the cell of a nonmetabolizable inducer are entirely and secondary, thena purely negative action of this kind would follow. This would mean, however, that the regulatory gene had no real function other than to cut off a normally quite unnecessary process, namely the formation of inducible enzyme. If it had no other function, then natural selection would seem to have done its work very badly, leaving two genes with no function but to frustrate one another.*"

PART III : Induction

Outline

Mill's Methods of Induction

- Mill's Methods are five methods of induction. They were described by philosopher John Stuart Mill in his 1843 book "*A System of Logic*."
- This is the first attempt to understand "causation" systematically.[2]

[2] Three of these methods, namely the methods of **agreement**, **difference** and **concomitant variation**, were first described by Avicenna, (Abu Ali Sina Balkhi), in his 1025 book "The Canon of Medicine." The remaining two methods, namely the method of **residues** and the joint method of **agreement and difference**, were first described by Mill.

Mill's Methods

1. Direct Method of agreement
2. Method of difference
3. Joint method of agreement and difference
4. Method of residues
5. Method of concomitant variations

Direct Method of Agreement

Direct Method of Agreement

If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon.

- For a property to be a necessary condition it must always be present if the effect is present.
- Since this is so, then we are interested in looking at cases where the effect is present and taking note of which properties, among those considered to be 'possible necessary conditions' are present and which are absent.
- Obviously, any properties which are absent when the effect is present cannot be necessary conditions for the effect.

Direct Method of Agreement

A	B	C	D	E	F	G	w	t	u	v	x	y	z
P	P	P	P	A	A	A	P	A	A	A	P	P	P
P	A	A	A	P	P	P	P	P	P	P	A	A	A

Therefore A is the cause, the effect, or part of the cause of w.

Method of Difference

Method of Difference

If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon.

Method of Difference

A	B	C	D	E	F	G	w	t	u	v	x	y	z
P	P	P	P	A	A	A	P	A	A	A	P	P	P
A	P	P	P	A	A	A	A	A	A	A	P	P	P

Therefore A is the cause, the effect, or part of the cause of w.

Joint Method of Agreement and Difference

Joint Method of Agreement and Difference

If two or more instances in which the phenomenon occurs have only one circumstance in common, while two or more instances in which it does not occur have nothing in common save the absence of that circumstance: the circumstance in which alone the two sets of instances differ, is the effect, or cause, or a necessary part of the cause, of the phenomenon.

- Also called simply the “joint method,” this principle simply represents the application of the methods of agreement and difference.

Joint Method of Agreement and Difference

A	B	C	D	E	F	G	w	t	u	v	x	y	z
P	P	P	A	A	A	A	A	A	A	A	P	P	P
P	A	A	P	P	A	A	P	A	A	A	P	P	A
A	P	P	A	A	A	A	A	A	A	A	A	P	P

Therefore A is the cause, the effect, or part of the cause of x.

Method of Residues

Method of Residues

Deduct from any phenomenon such part as is known by previous inductions to be the effect of certain antecedents, and the residue of the phenomenon is the effect of the remaining antecedents.

- If a range of factors are believed to cause a range of phenomena, and we have matched all the factors, except one, with all the phenomena, except one, then the remaining phenomenon can be attributed to the remaining factor.

Method of Residues

A	B	C	D	E	F	G	w	t	u	v	x	y	z
P	P	P	A	A	A	A	A	A	A	A	P	P	P

B is known to be the cause of y
C is known to be the cause of z

Therefore A is the cause or effect of x.

Method of Concomitant Variations

Method of Concomitant Variations

Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation.

- If across a range of circumstances leading to a phenomenon, some property of the phenomenon varies in tandem with some factor existing in the circumstances, then the phenomenon can be attributed to that factor.

Method of Concomitant Variations

A	B	C	D	E	F	G	w	t	u	v	x	y	z
P	P	P	A	A	A	A	A	A	A	A	P	P	P
↑	P	P	A	A	A	A	A	A	A	A	↑	P	P

Therefore A and x are causally connected

The Structure of Complex Statements

- **Logical Connectives**: Propositional Connectives such as 'not' ($\neg$), 'and' ($\wedge$) and 'or' ($\vee$) can be used to create more complex statements.
- Mill's methods can be extended in a natural manner to include complex statements in 'propositional logic.'
- Statements can be verified (in a model) using truth tables. **Model Checkers** can perform these verifications inductively starting with atomic propositions.

Observations

- From observing event E n times, we cannot predict that event E will occur also on the nth time. There are many unmodelled disturbances.
- E.g., if in the first 10 tosses a coin displays only heads, we may not still conclude that from 11 onward, the coin toss will result only in heads.
- *Belief Revision*.

The Goodman Paradox

- Inductive reasoning gets into trouble because of the linguistic constructs used in defining various situations.
- An object is *grue* (resp. *bleen*) iff it is green (resp. blue) before 2011 and blue (resp. green) after 2011.
- Define an object to be *green* if it is grue before 2011 and bleen after 2011. Similarly, define *blue*.
- From the observations that algae have been observed to be grue (without exception), conclude that in the future algae will be observed to be grue.

[End of Lecture #1]