How Science Thinks:
The Science and Engineering of Science and Engineering

What is engineering?

Engineering is a search for artifacts that serve particular functions.

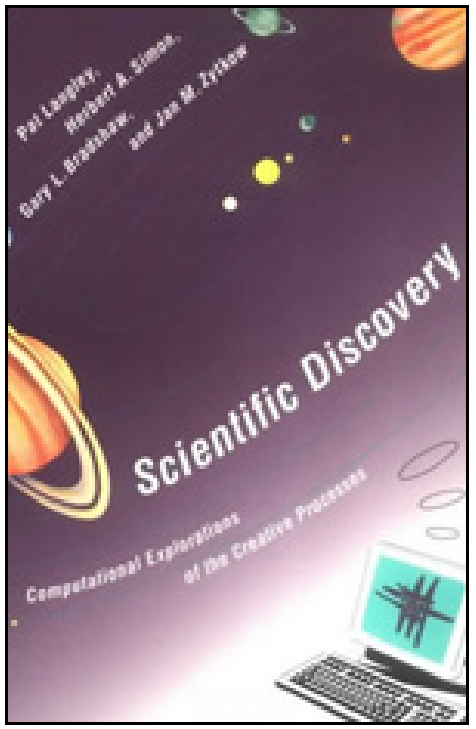
What is science?

Science is a search for artifacts called models (or theories)
that can explain and predict phenomena.

Wait! So is science just a type of engineering?

Um, yeah, I guess so...

Okay smart ass, so what's search!?



How Science Thinks:
The Science and Engineering of Science and Engineering

What is Search?

(It's *not* what Google does; that's more like "lookup").

> *Search is the exploration of a space of possibilities for one or more that satisfy a particular goal.*

Gimme some examples!

Searching for the right play in a game like football or Go.

Searching for your cell phone.

Searching for a place to eat tonight.

Searching for an way to detect gravity waves.

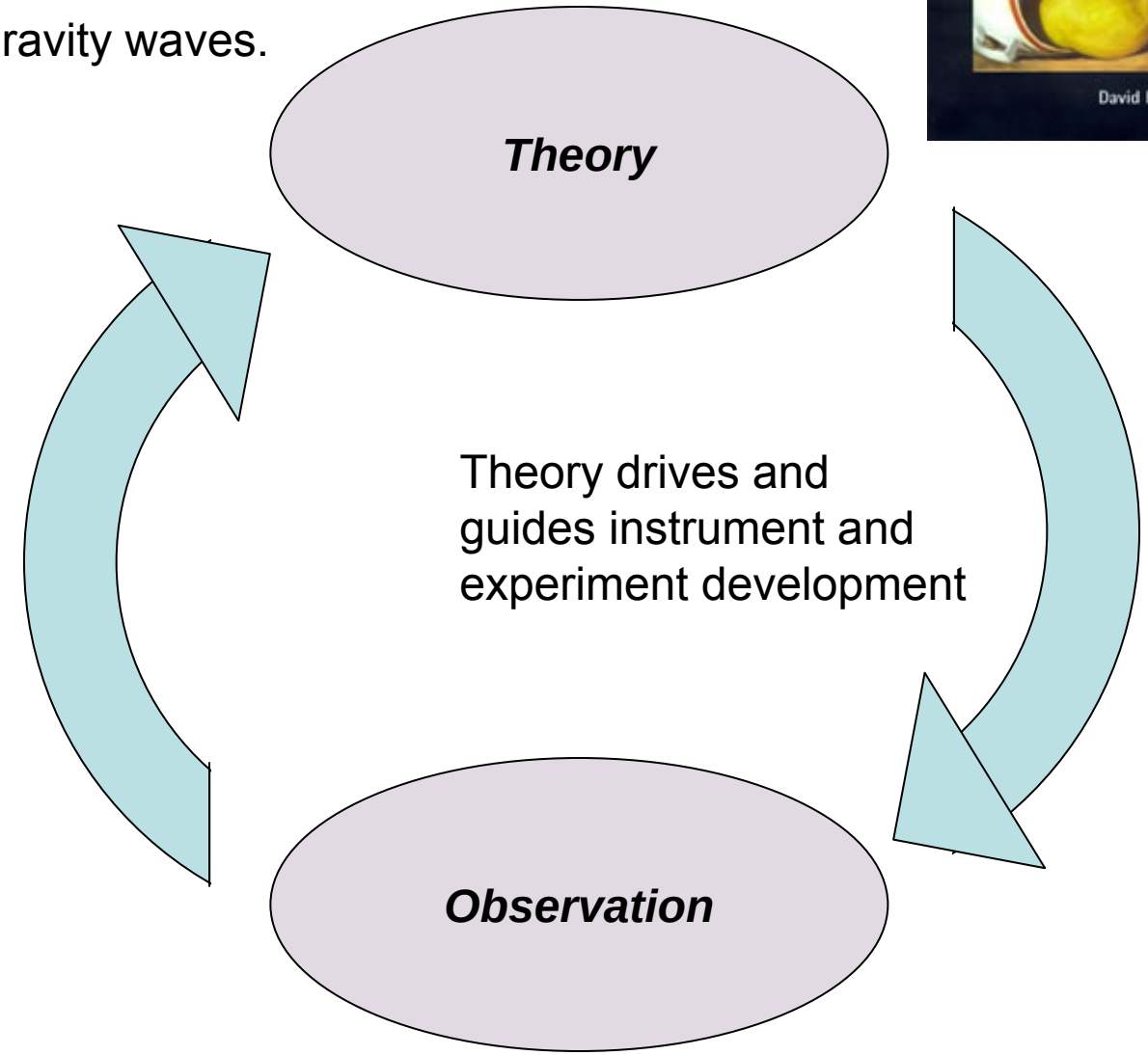
Searching for a theory of gravity waves.

Searching for a cure for cancer.

How Science Thinks:
The Science and Engineering of Science and Engineering

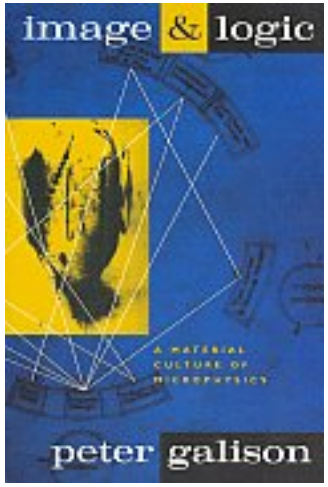
Searching for a way to detect gravity waves.

Searching for a theory of gravity waves.



Theory drives and guides instrument and experiment development

Observations drive and guide theory modification



How Science Thinks:

The Science and Engineering of Science and Engineering

Searching for an instrument that will detect gravity waves.

Searching for a theory of gravity waves.

How big is the search space?

$E=mc^2$

$E=mc^3$

$E=mc^4$

$E=mc^5$

$E=mc^6$

$E=mc^7$

$E=mc^8$

$E=...$

For every right answer in science,
There is an infintitude(-1) of wrong ones!

Historical Theory↔Experiment SeeSawing of the Gyromagnetic Ratio “g-factor”

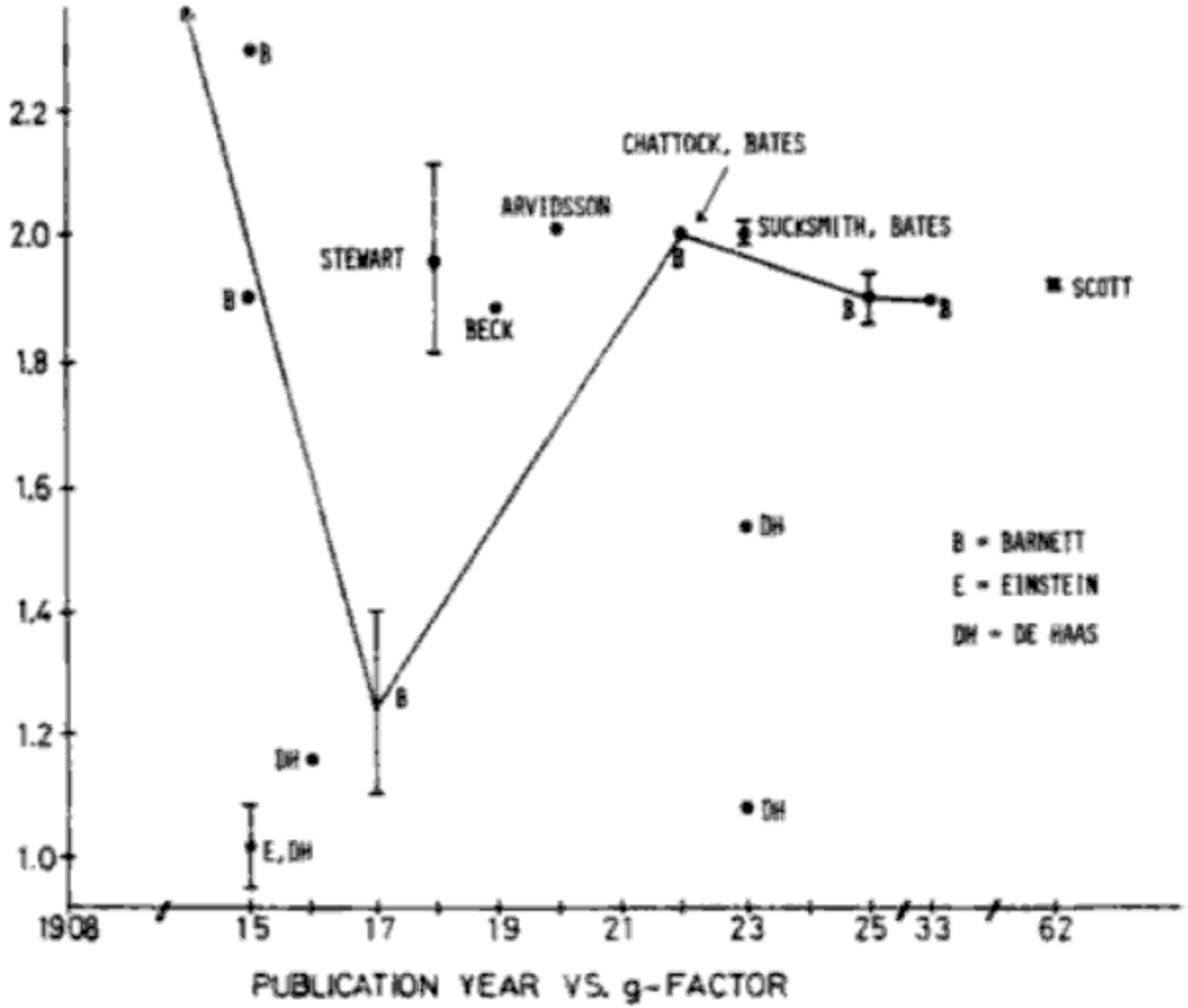
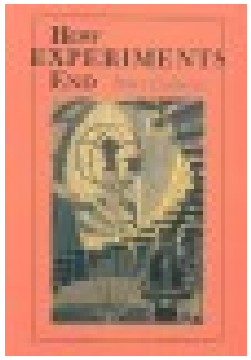


FIGURE 2.15. Publication year versus g-factor. The solid line traces Barnett's results as a function of time. Einstein's theoretical prediction was that $g = 1$; later experiment and theory both gave $g \sim 2$.



Galison (1987) How Experiments End. Chicago U. Press

How Science Thinks:
The Science and Engineering of Science and Engineering

For every right answer in science,
There is an infintitude(-1) of wrong ones!

HOW COULD THIS EVER WORK!?

1. Close is often good enough, or at least guides you to the right answer.

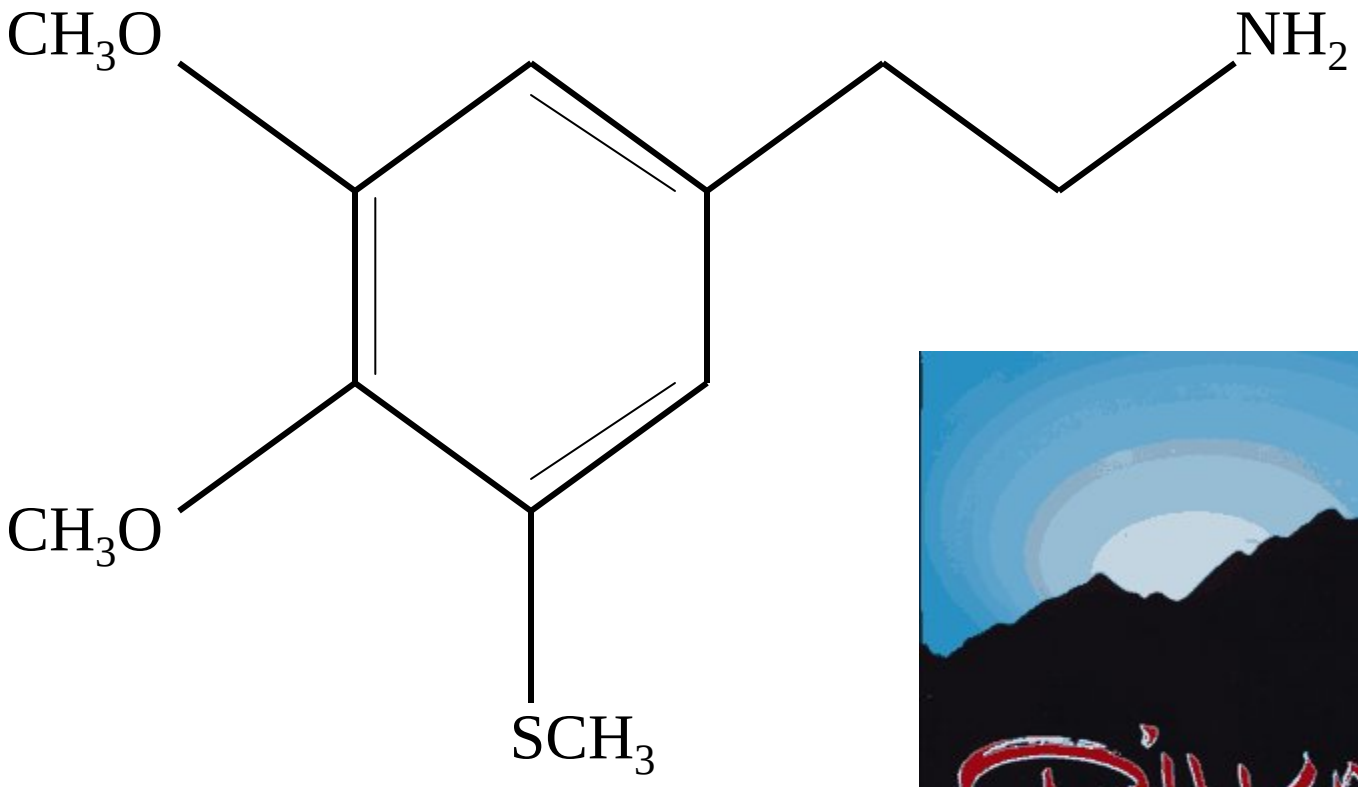
2. Theory (model) guidance reduces the search space by huge orders!

3. We've been really *really* lucky ... so far, anyway!

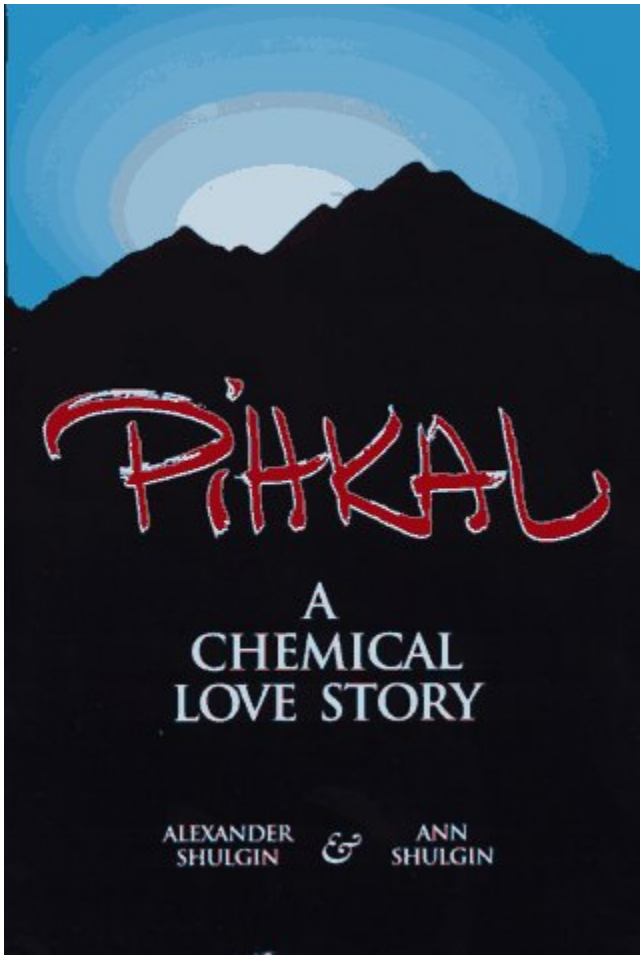
4. *You are not alone!* (>15Million abstract in pubmed alone!)

Pine trees seem a good place
To start. Notwithstanding this table
Of pine, unfinished, unruled,
The pulp upon which we reveal
The unnerved thought.
How casual we are at discarding
Our feelings, a rubble we
Leave behind for the living.
Who among us can absorb
The spiritual load we see as
What others carry.

Alexander Shulgin
PIHKAL 1991



3,4-dimethoxy-5-methylthiophen-ethylamine (60-100mg orally)



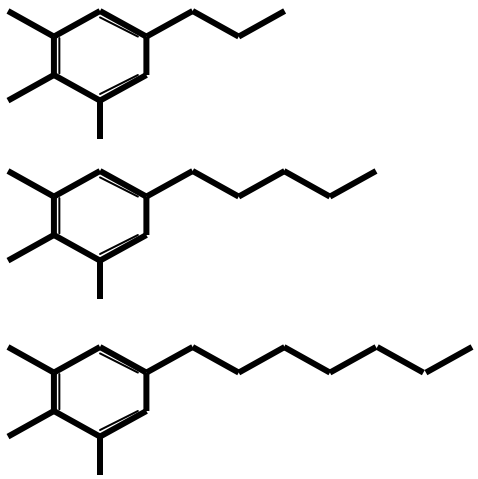
The Science and Engineering of "Drug Discovery"

How big is the search space?

Searching for a theory of gravity waves.

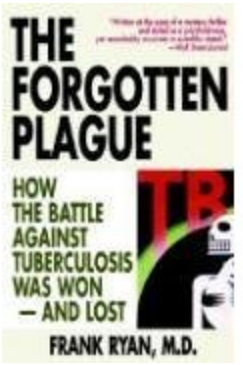
$E=mc^2$
$E=mc^3$
$E=mc^4$
$E=mc^5$
$E=mc^6$
$E=mc^7$
$E=mc^8$
$E=...$

Searching for a cure for cancer.

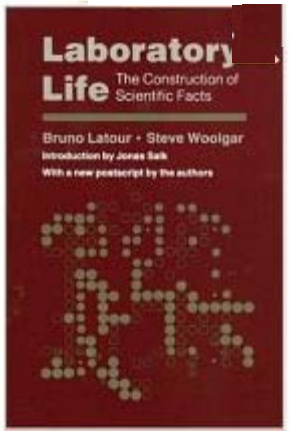


For every right answer in science,
There is an infintitude(-1) of wrong ones!

The Science and Engineering of "Drug Discovery"



"Albert [Schatz] hunted for new strains of actinomyces in soil, in manure heaps, in drains, even from the culture plates that were being thrown away by colleagues working on other unrelated projects, indeed anywhere in the world that his imagination would take him—this was Albert's entire life." (p. 215)



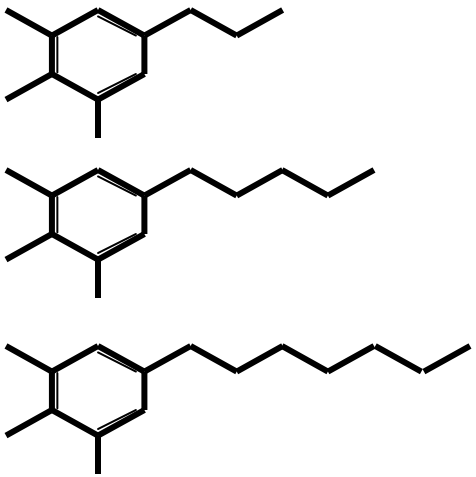
"It was salt mine, where, in order to pull a practical antibiotic producer out of Mother Nature, we literally have to work our asses off. The failure rate is about 99.99 per cent" (p. 218).

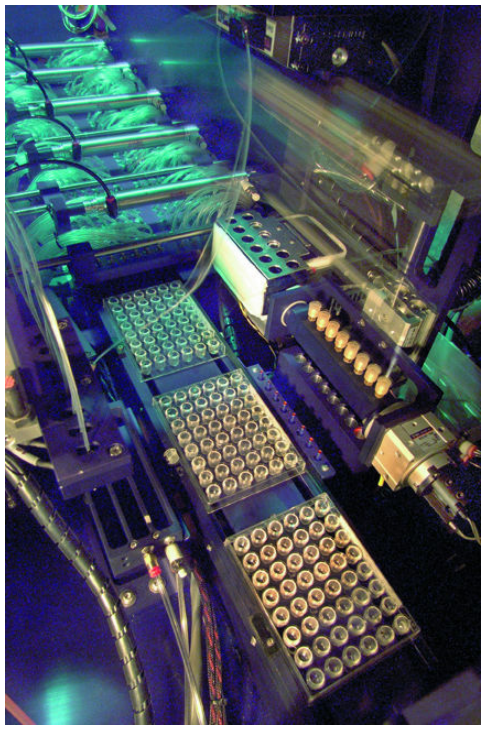
"Using techniques that seem closer to gardening than the intellectual exercise of science, Rene [Dubos] trowelled soil into pots, searched in farmers' fields, manure heaps, lawns and hedges, altered growing conditions, added and subtracted chemicals. (p. 65)

The Science and Engineering of "Drug Discovery"

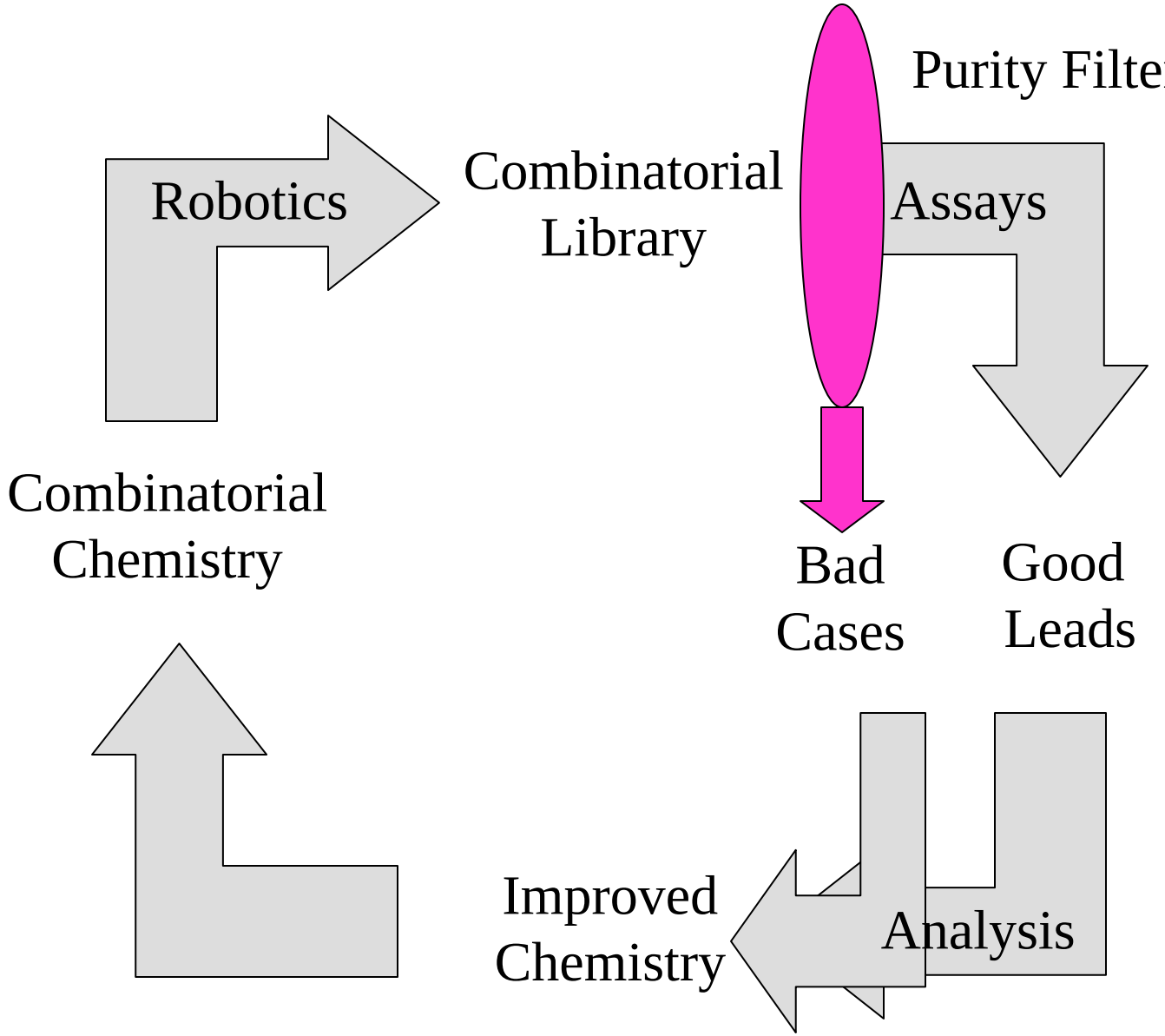
Afferent helps medicinal chemists do lead discovery.

Drives the (robotic) synthesis of combinatorial reactions.
Closes the synthetic/analytic loop on drug (lead) discovery.
Gives scientists direct control over the search process.

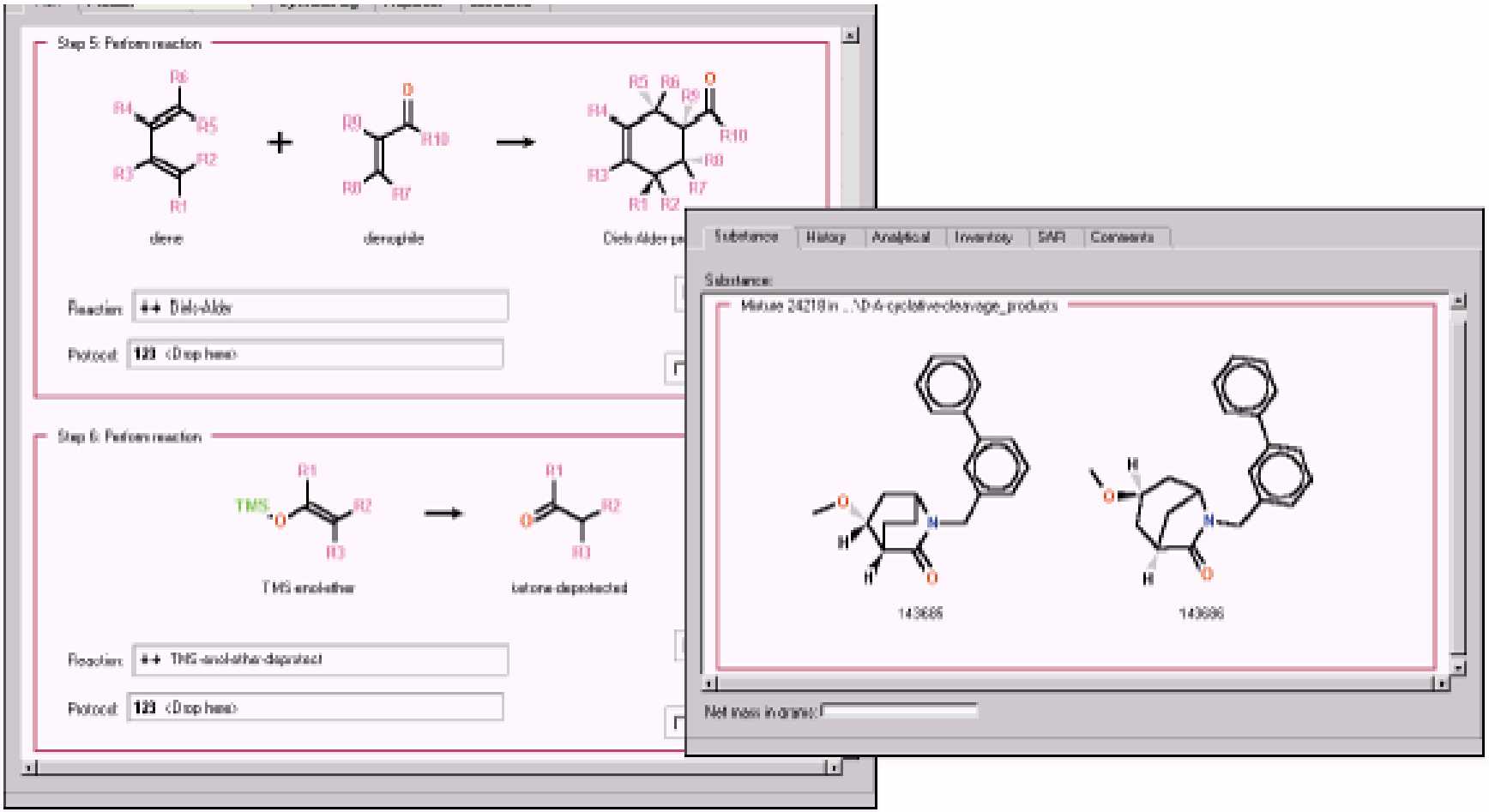




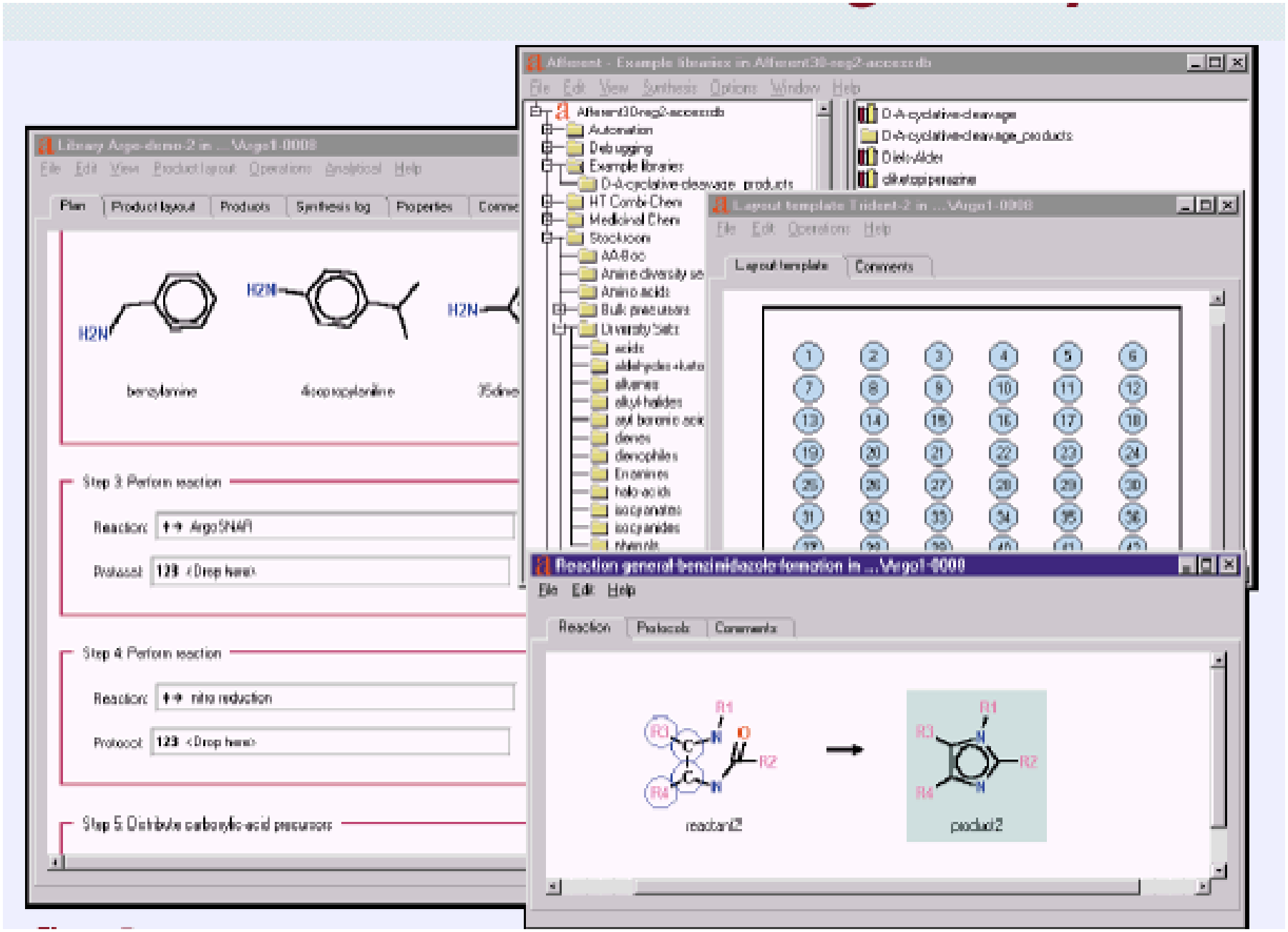
Combinatorial Drug Discovery -- Closing the Loop



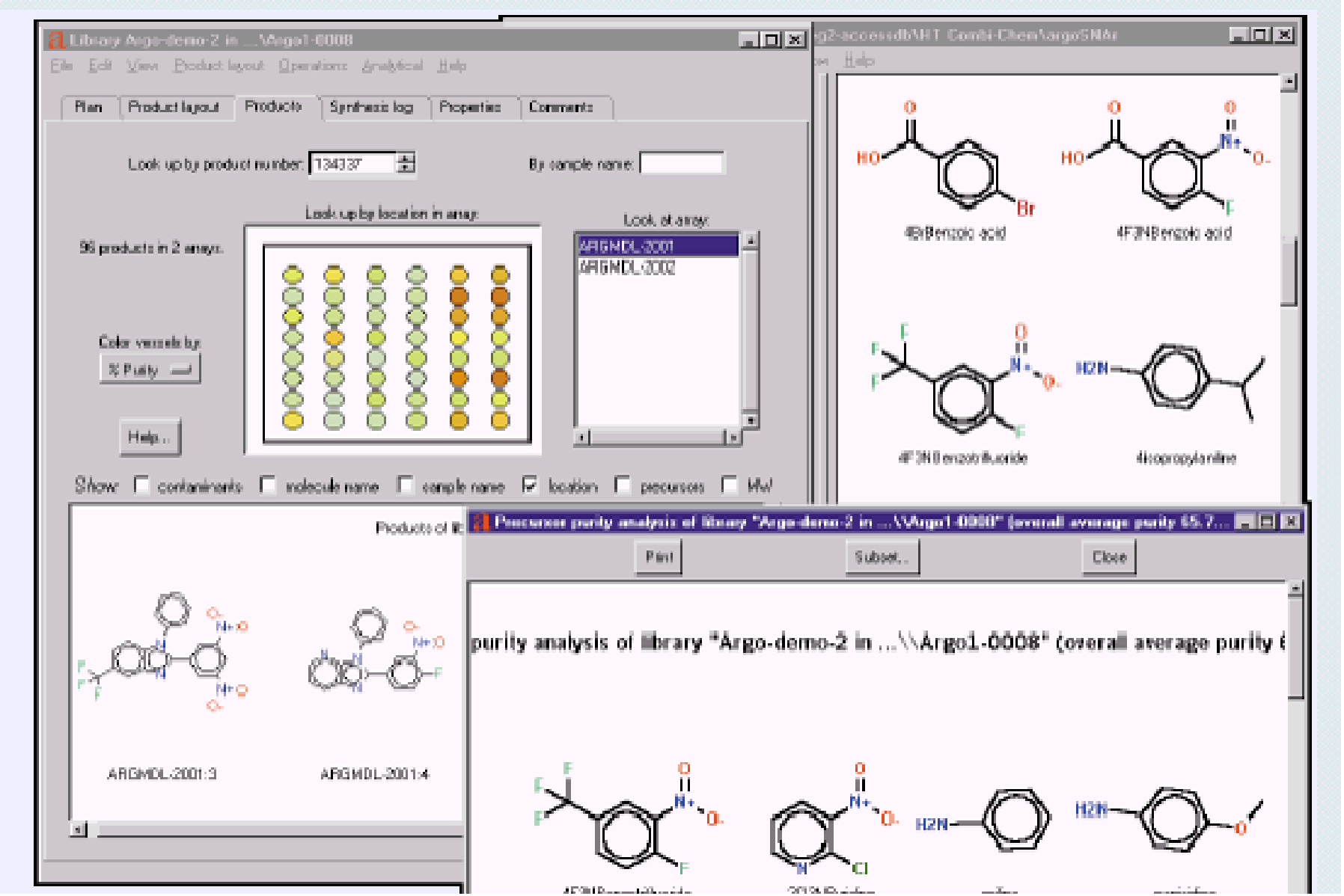
Chemists “Teach” Afferent Organic Chemistry



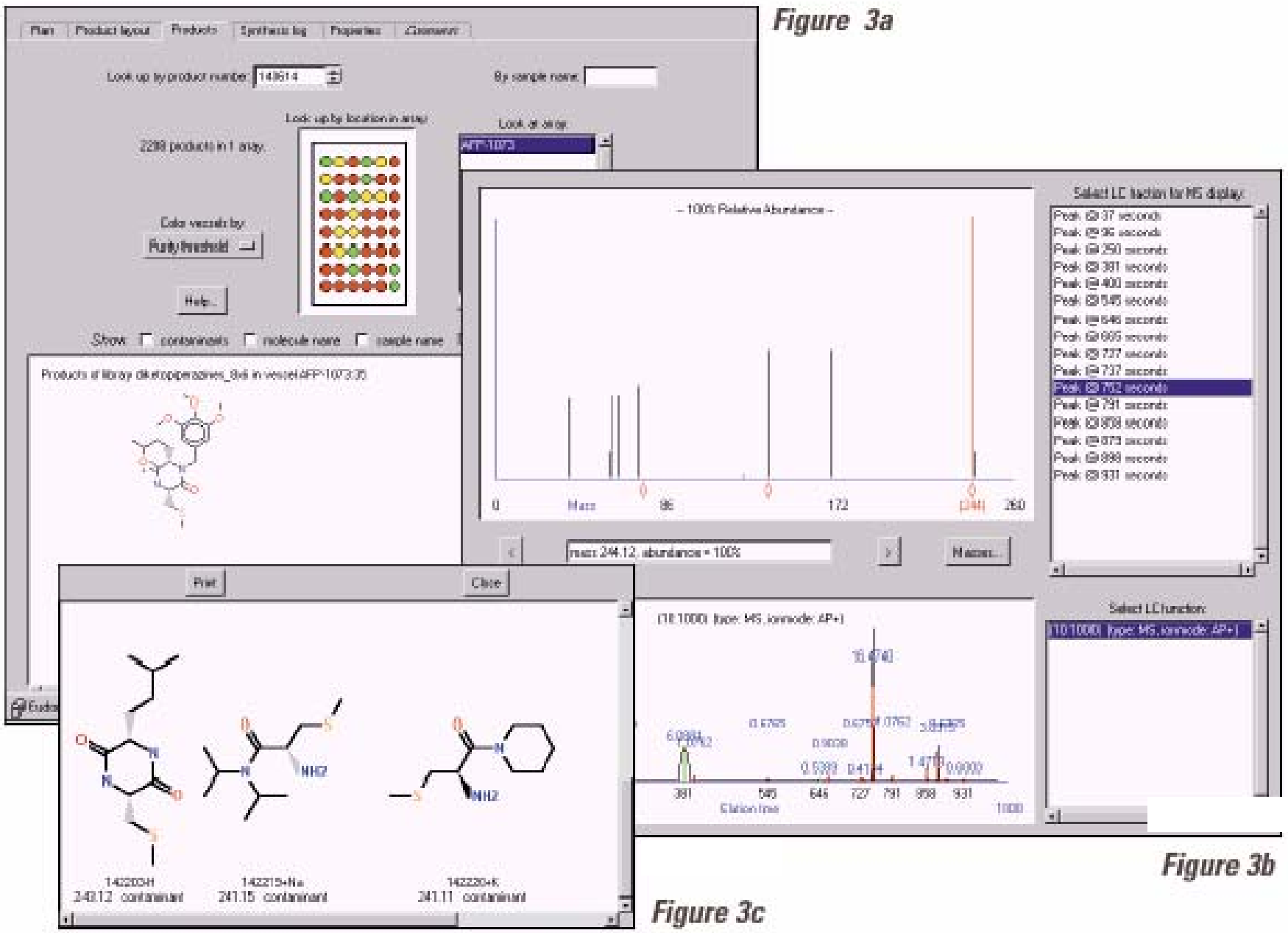
Afferent Runs Chemical Robots to Do the Reactions



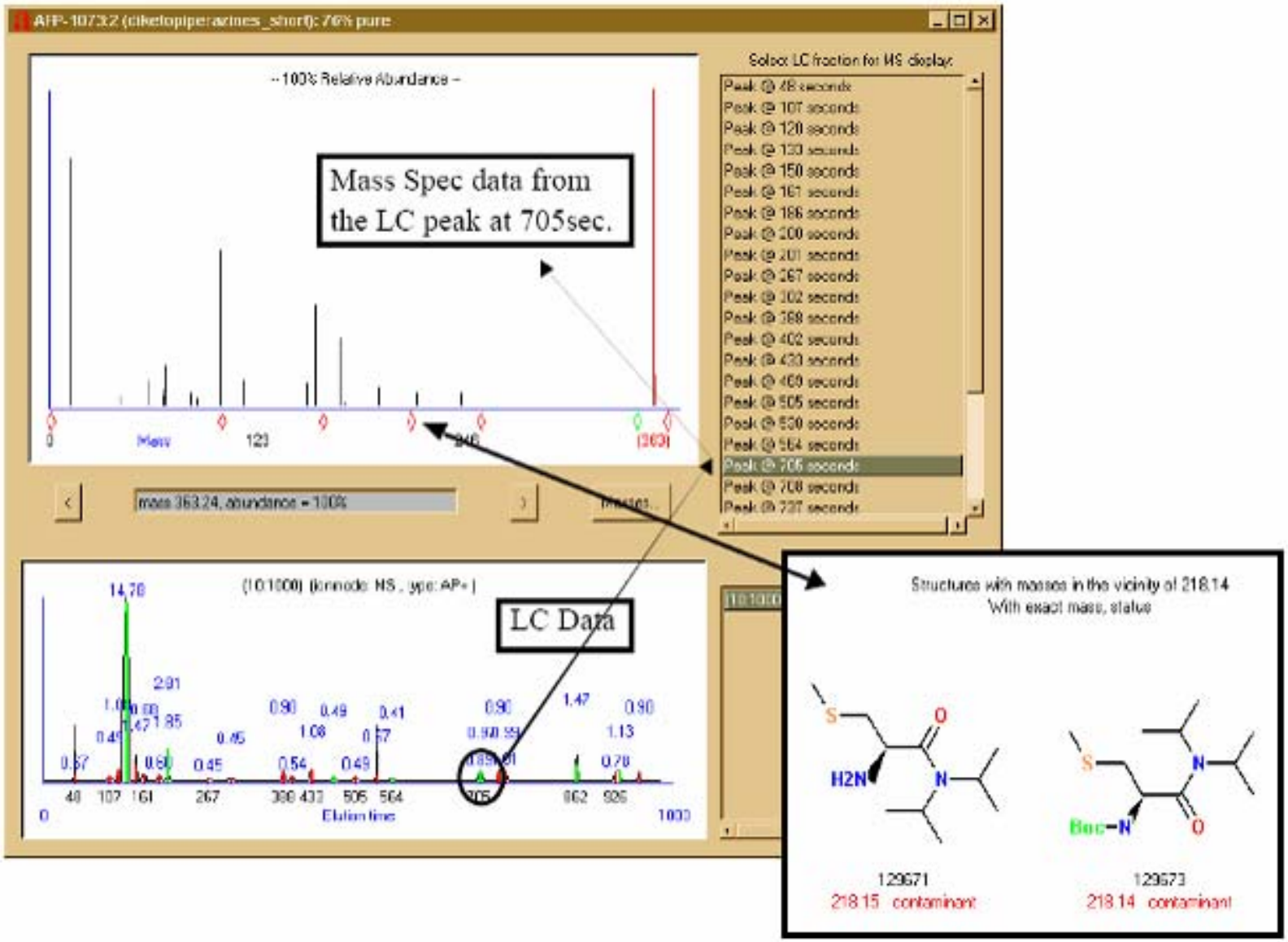
Afferent Simulates Combinatorial Chemistry



Afferent can “see” both successes and failures in mass spec data.



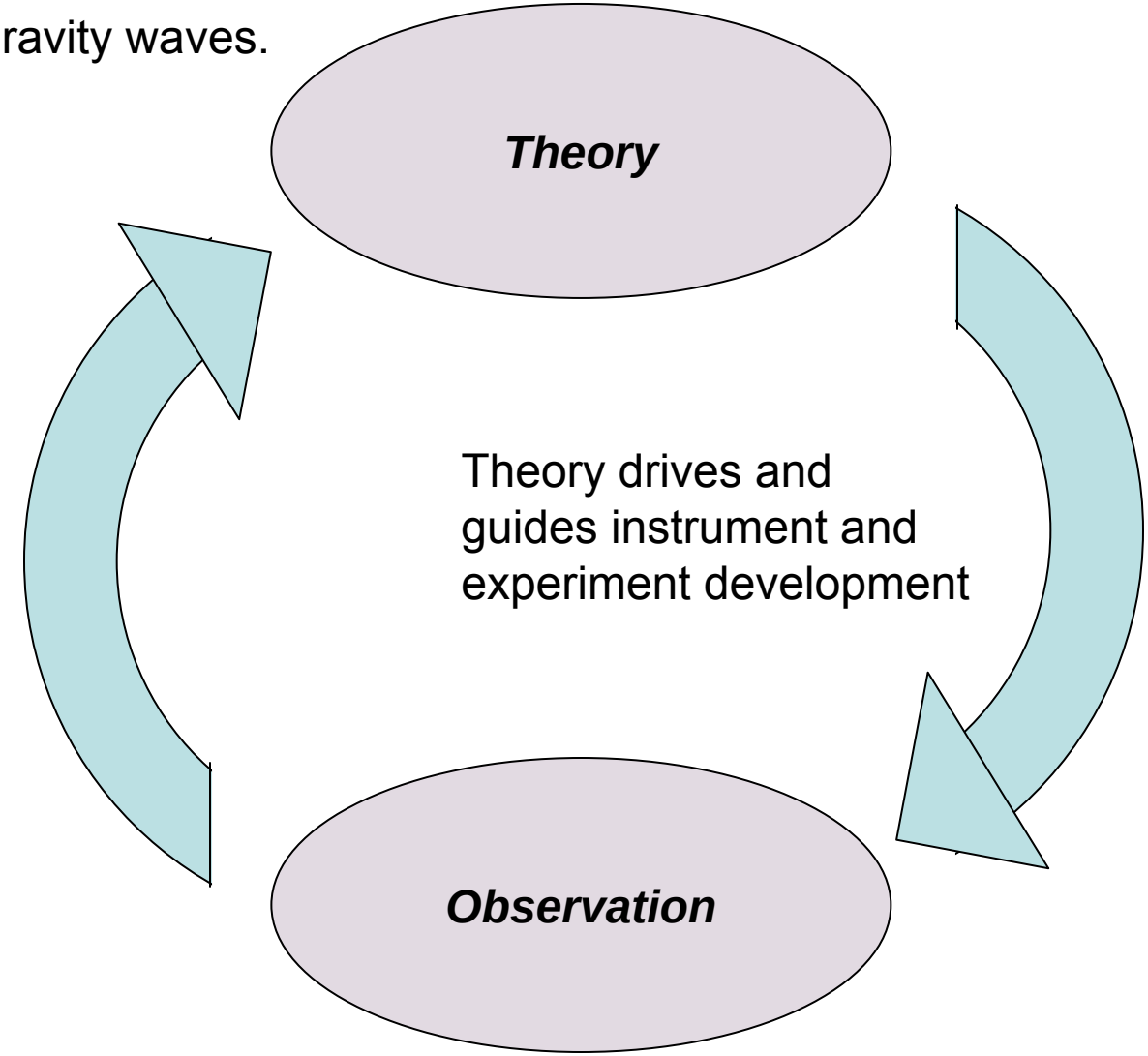
Afferent can make “educated guesses” about what might have gone wrong



How Science Thinks:
The Science and Engineering of Science and Engineering

Searching for an instrument that will detect gravity waves.

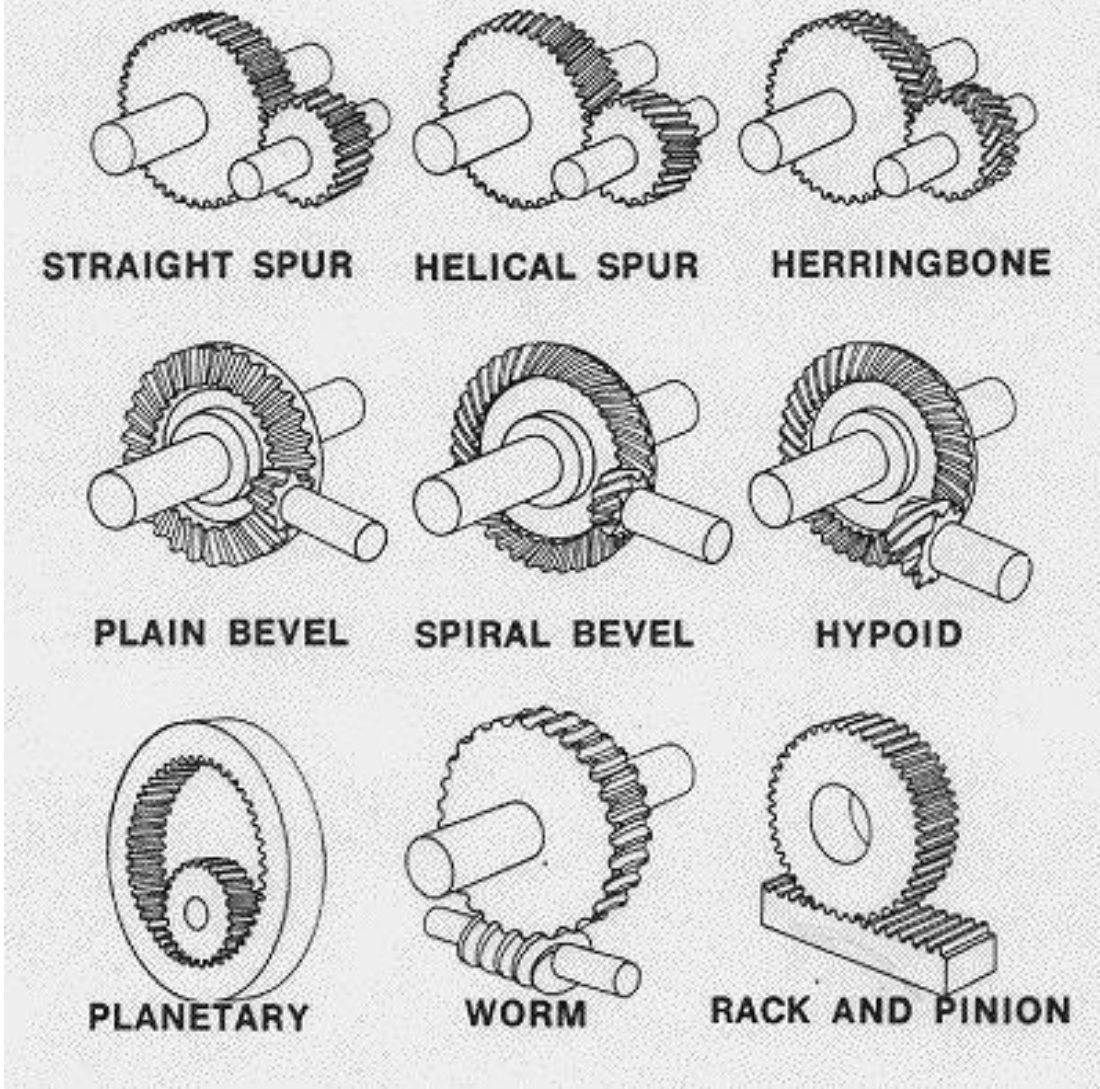
Searching for a theory of gravity waves.



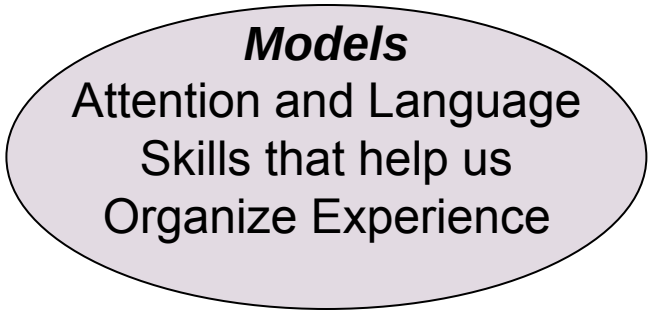
Explanation is the main function of theories (models)

http://files.turbosquid.com/PreModel/Content_on_8_29_2002_07_36_31/gears03.jpgDA4233CA-5F01-47A6-9BDFCDF1F8087F89.jpgLarge.jpg

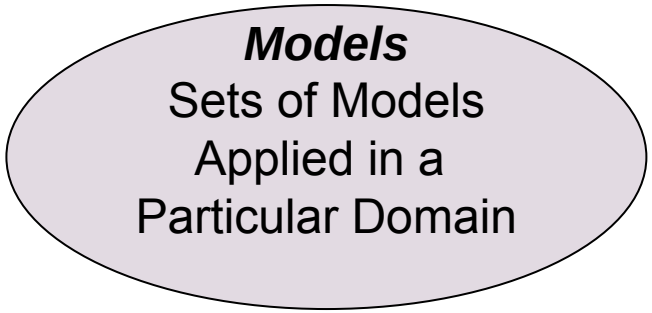
TYPES OF GEARS



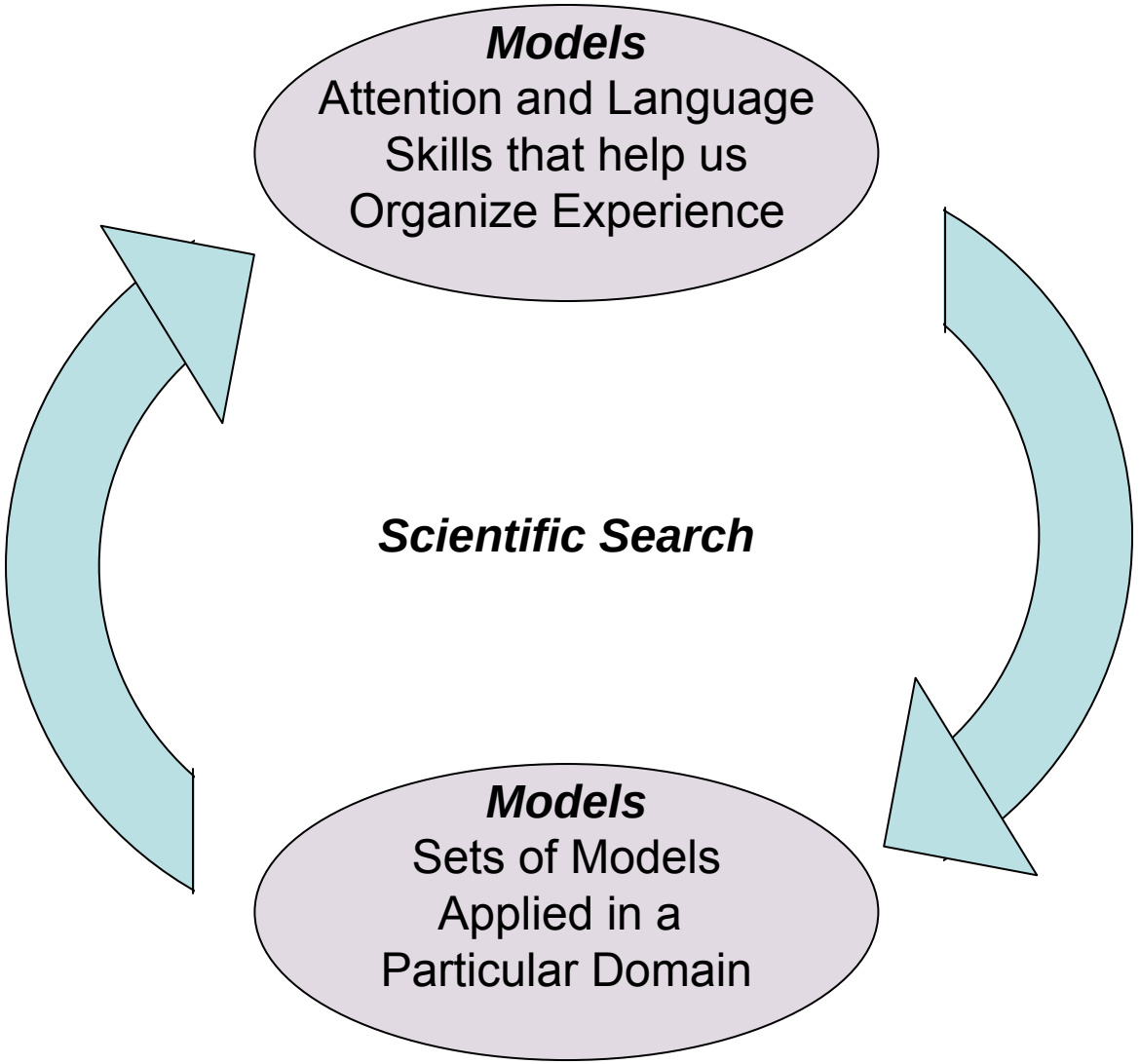
Explanation is the main function of theories (models)



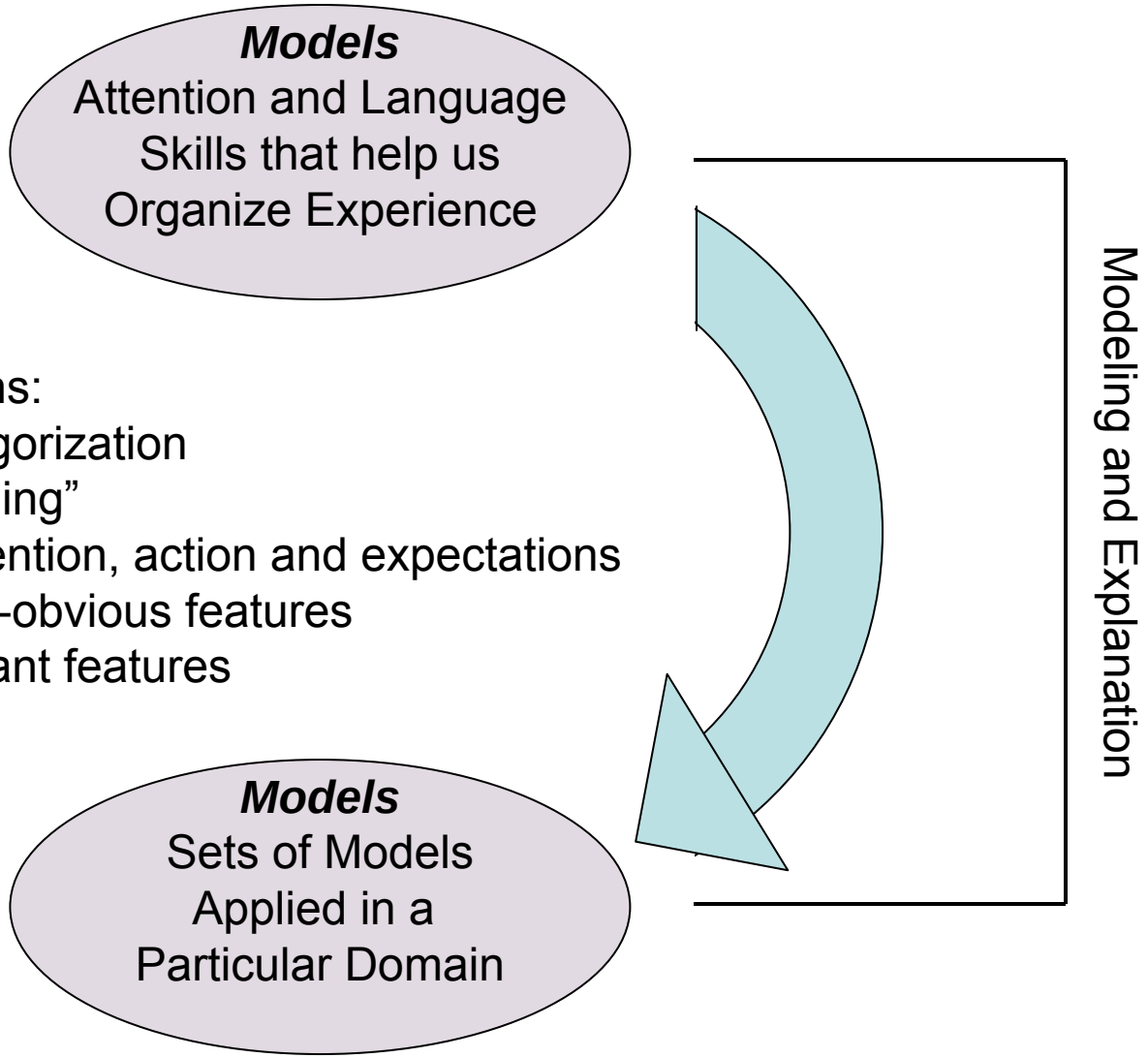
Model Application is a Cognitive Process through which we organize experience.
Explanation is the most obvious (public) features of this process.



Explanation is the main function of theories (models)



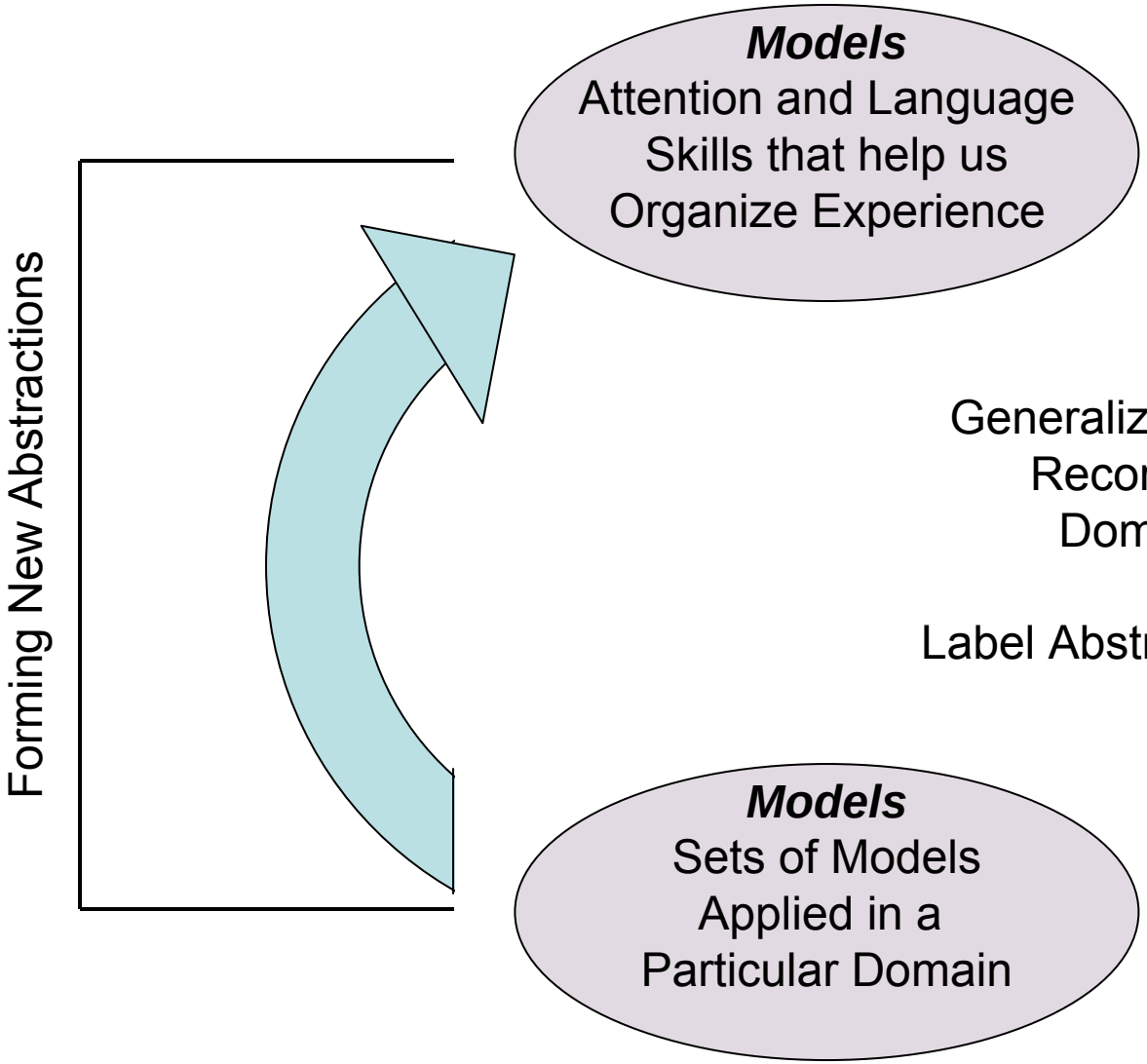
Explanation is the main function of theories (models)



Forming Explanations:

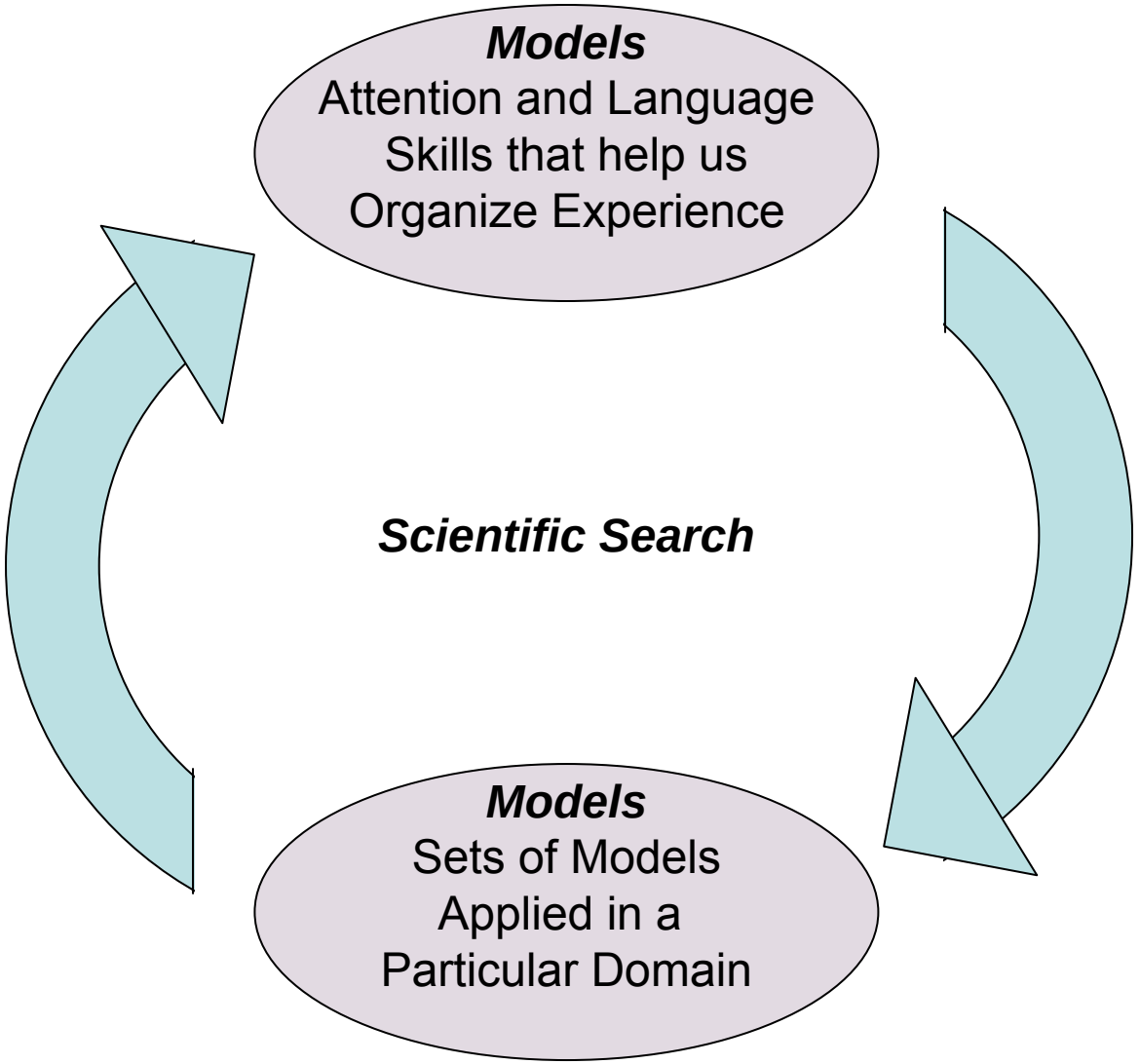
- Labeling and categorization
- "Conceptual Blending"
- Sequencing of attention, action and expectations
- "Discovery" of non-obvious features
- Focusing on relevant features

Explanation is the main function of theories (models)

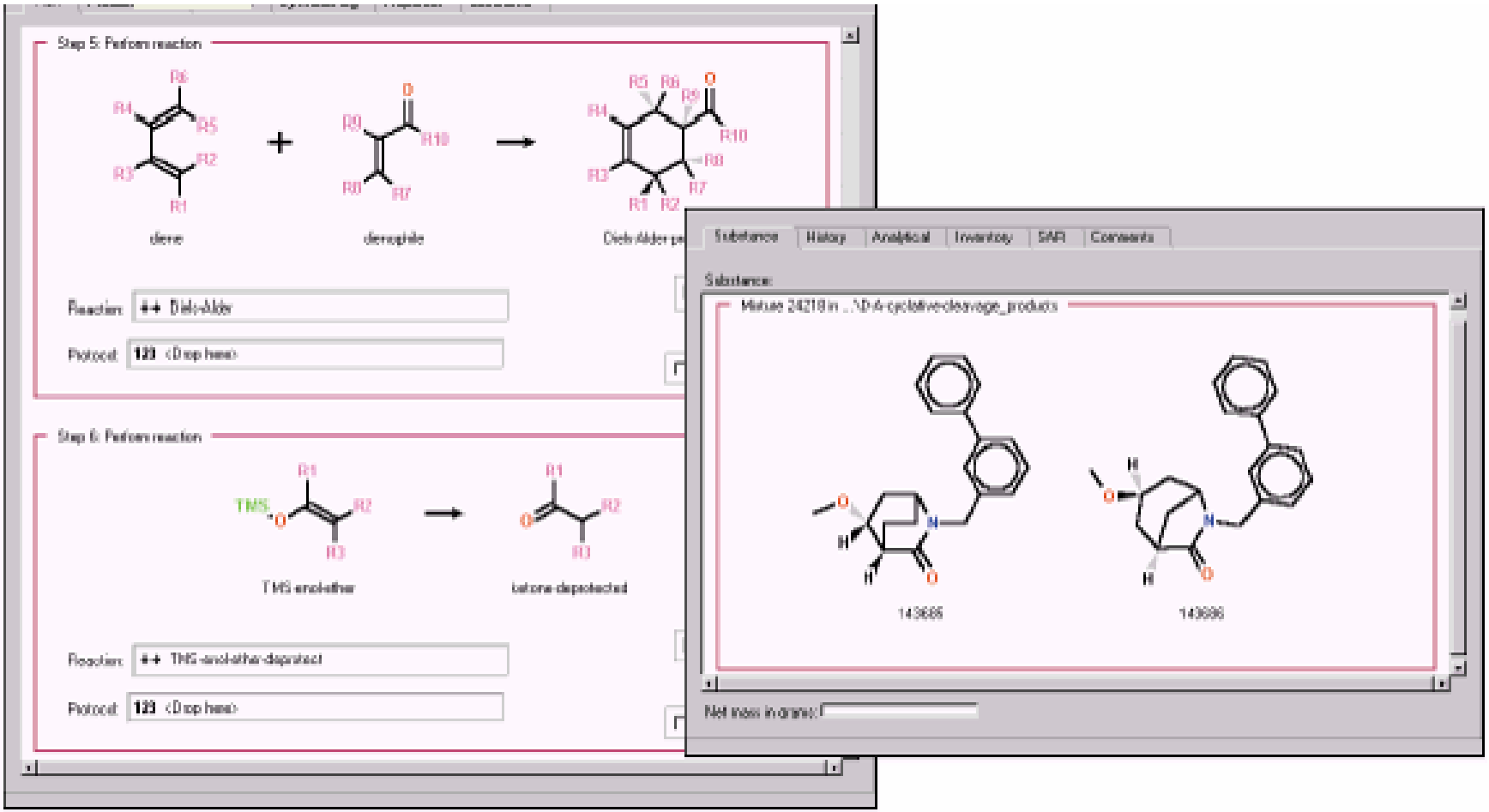


Generalization to New Models
Records for Later Analogy
Domain Characterization
Repetition/Practice
Label Abstraction (“Gene”, etc.)

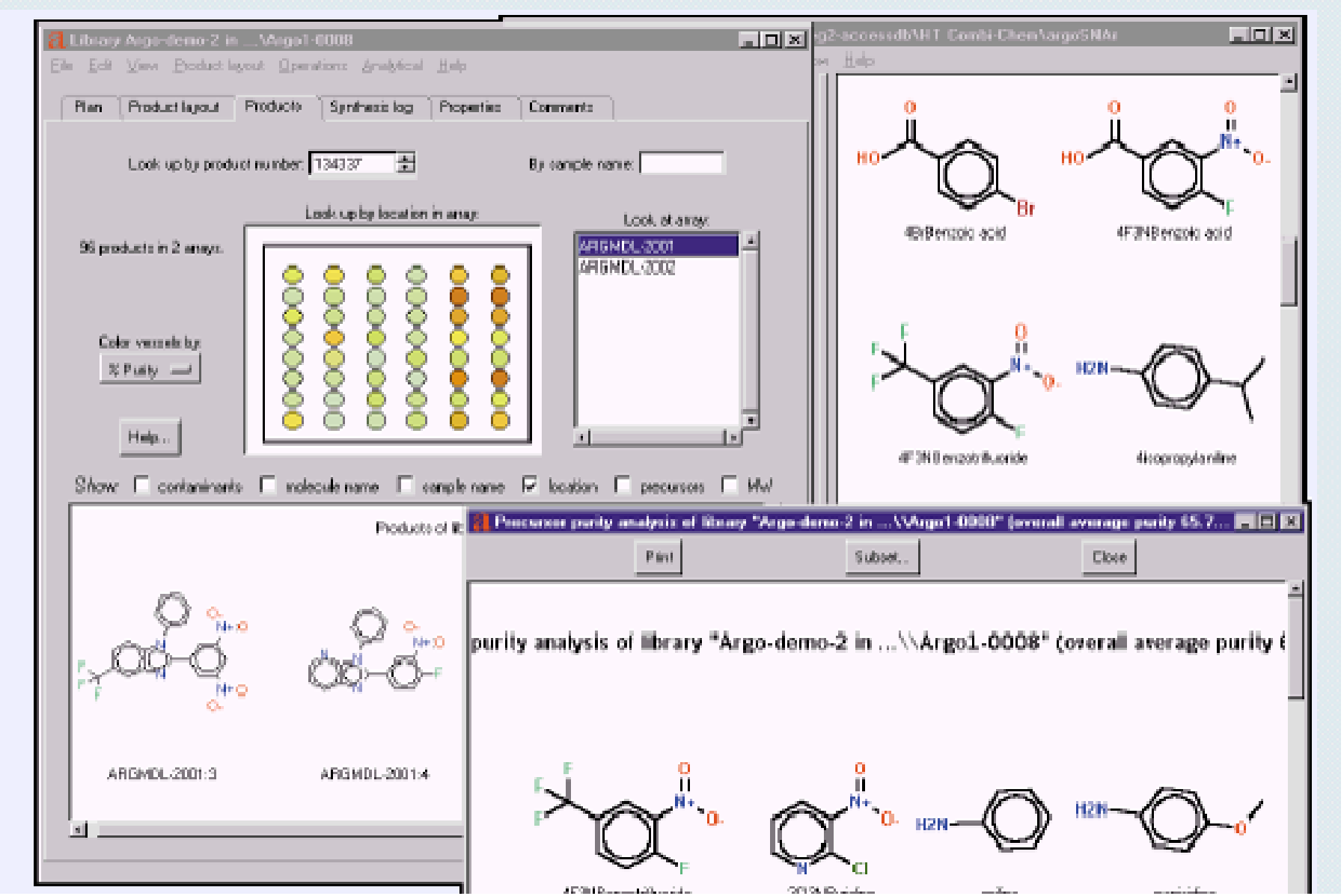
Explanation is the main function of theories (models)



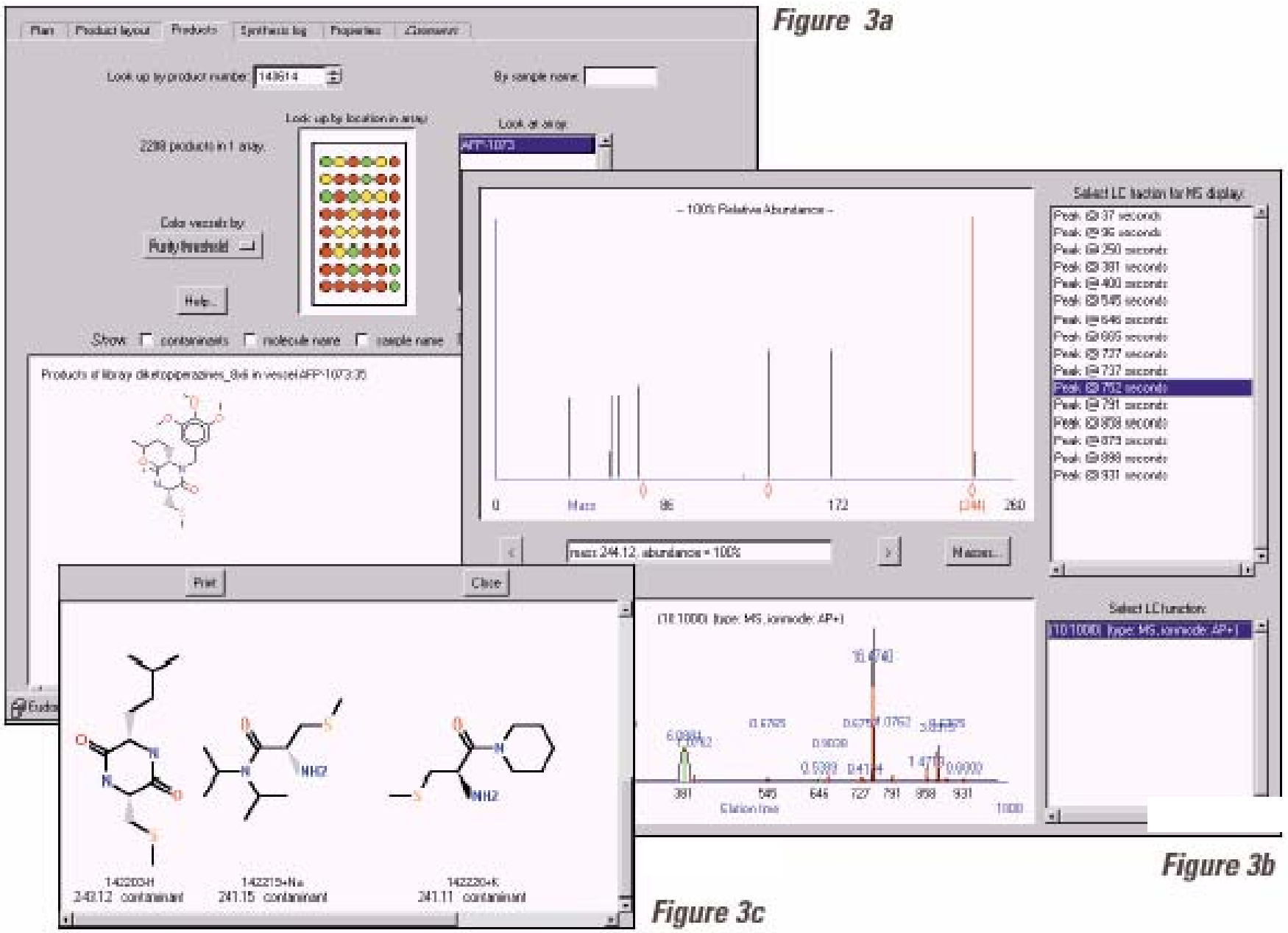
Chemists “Teach” Afferent Organic Chemistry



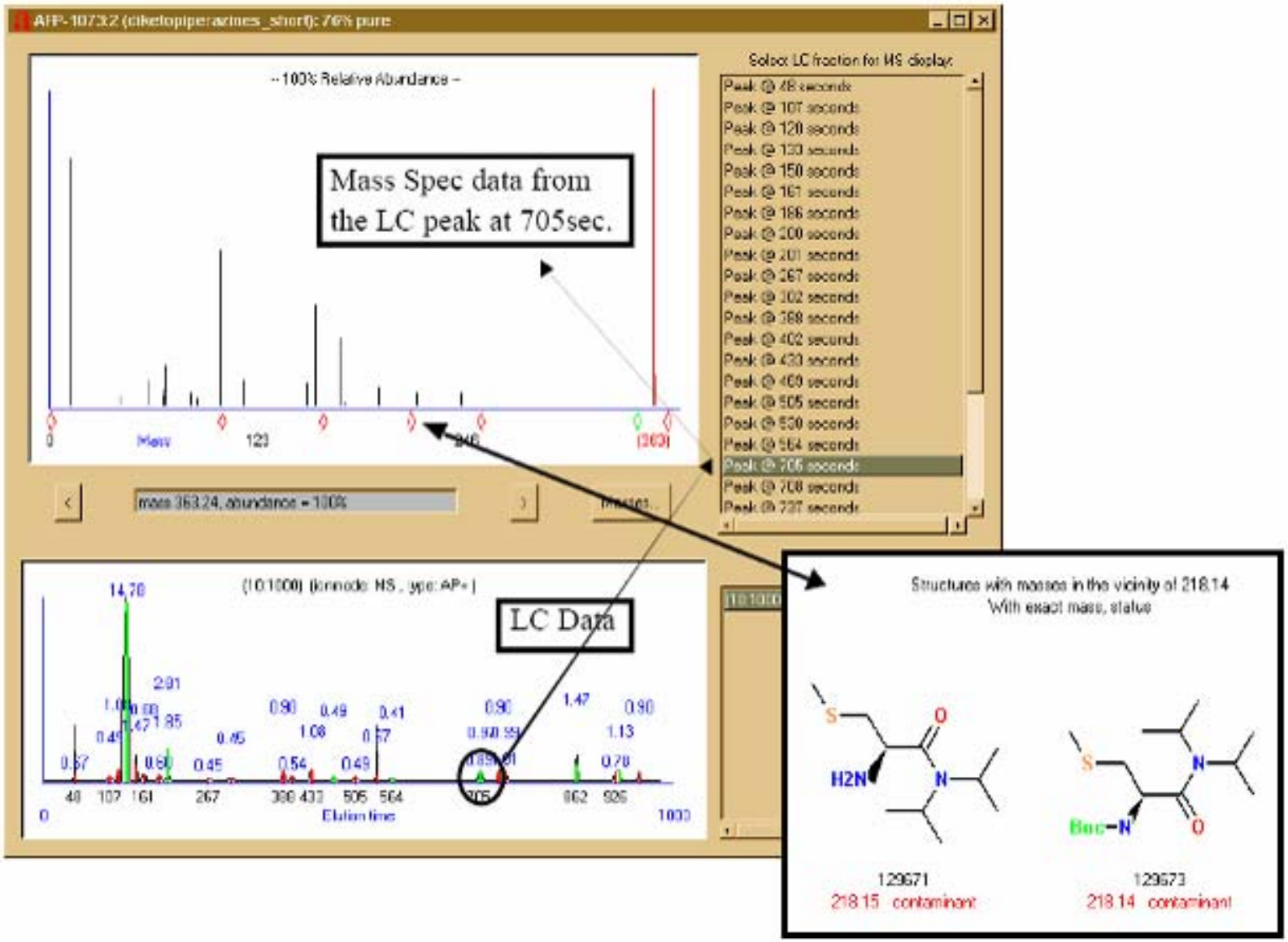
Afferent Simulates Combinatorial Chemistry



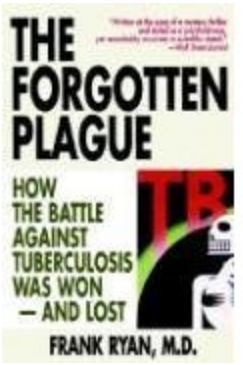
Afferent can “see” both successes and failures in mass spec data.



Afferent can make “educated guesses” about what might have gone wrong



The Science and Engineering of "Drug Discovery"

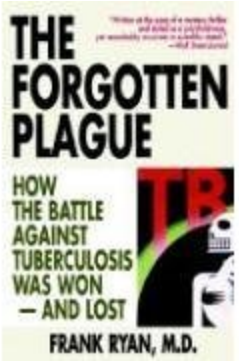


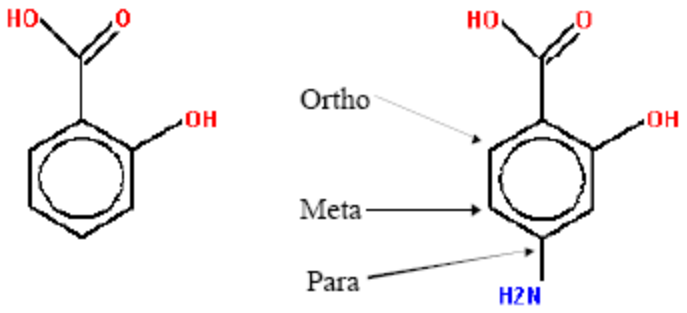
"Albert [Schatz] hunted for new strains of actinomyces in soil, in manure heaps, in drains, even from the culture plates that were being thrown away by colleagues working on other unrelated projects, indeed anywhere in the world that his imagination would take him—this was Albert's entire life." (p. 215)

"It was salt mine, where, in order to pull a practical antibiotic producer out of Mother Nature, we literally have to work our asses off. The failure rate is about 99.99 per cent" (p. 218).

"Using techniques that seem closer to gardening than the intellectual exercise of science, Rene [Dubos] trowelled soil into pots, searched in farmers' fields, manure heaps, lawns and hedges, altered growing conditions, added and subtracted chemicals. (p. 65)

The Science and Engineering of "Drug Discovery"





Salicylic Acid (2-hydroxybenzoic acid)

PAS (p-Aminosalicylic Acid;
4-Amino-2-hydroxybenzoic acid)

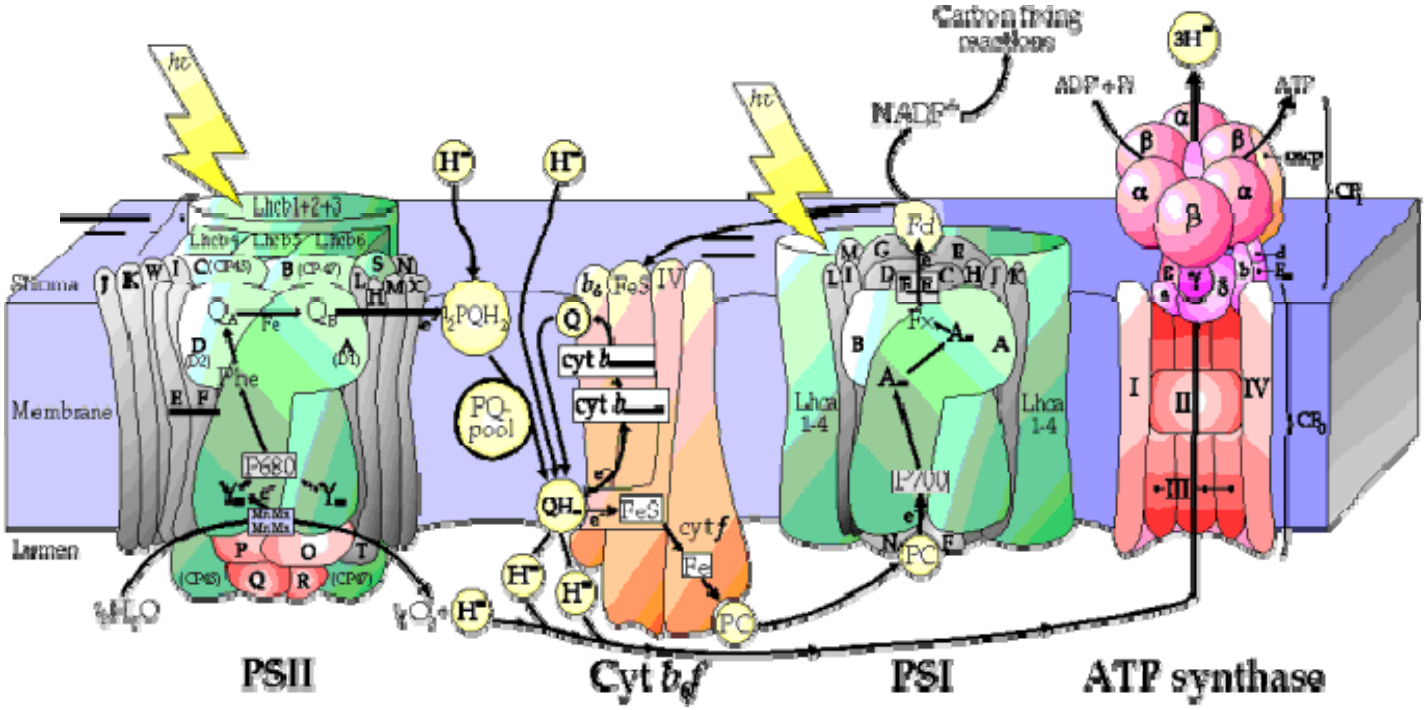
Figure 1. PAS, para-amino salicylic acid (right) was the drug discovered by Jorgen Lehmann in 1943. It is built upon a salicylic acid core (left), which is the same core upon which aspirin is built, by adding an amino group (H2N) in the "para" position (6 0'clock) of the central ring. O=oxygen, N=nitrogen, H=hydrogen. Each kink in the ring represents the location of a carbon atom, and there are hydrogen atoms attached in various places. Hydrogen atoms are usually not shown in organic representation because organic chemists can tell right away how many there are and where they must go. The OH and H2N groups are organic "idioms," recognized by chemists. Lines connecting atoms represent single or double bonds.

"[It] was a deduction so brilliant that [Jorgen Lehmann's] fellow doctors and scientists would refuse to believe it. How could Lehmann have possibly picked out this single chemical derivative of aspirin as the one to test before a single experiment had been performed?" (p. 242)

Simulation: What does this model predict?

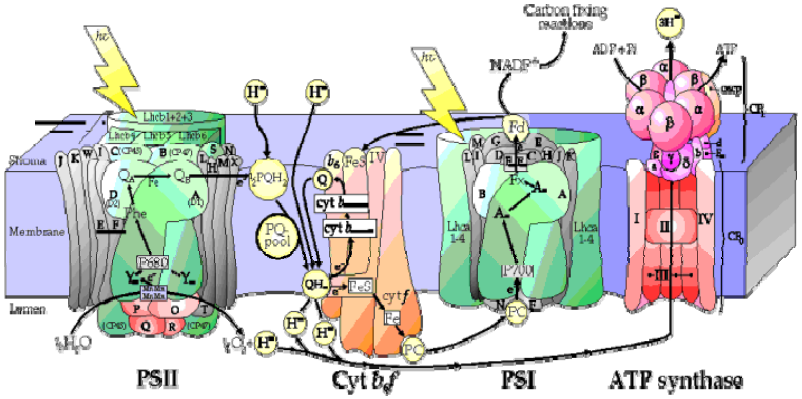
Explanation: How does it make these predictions?

Model Identification: What models fit this data?



Explanation by Pathway Tracing

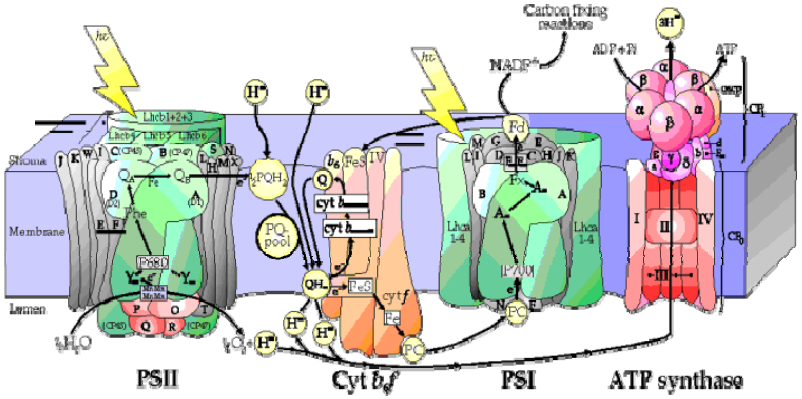
```
(photosynthesis isa process with
  inputs (chloroplast-inside.water everywhere.light chloroplast-outside.nadph+
          chloroplast-outside.adp chloroplast-outside.pi)
  outputs (chloroplast-outside.atp chloroplast-outside.nadph everywhere.o2)
  implemented-by photosystem)

(photosystem composition (psii antenna-array atpase pq-pool))

(light-absorption isa process with
  inputs (everywhere.light)
  outputs (chlorophyll.energy)
  function absorption
  implemented-by chlorophyll)

(light-energy-concentration isa process with
  outputs psii.energy
  driver chlorophyll.energy
  function concentration
  implemented-by antenna-array)
```




```
(psii-water-breakdown isa process with
  inputs (chloroplast-inside.water)
  driver psii.energy
  outputs (psii.e- psii.e- chloroplast-inside.h+ chloroplast-inside.o2)
  function molecular-splitting
  implemented-by psii)

(psii-pq-reduction isa process with
  inputs (psii.e- chloroplast-membrane.h+ chloroplast-membrane.plastoquinone)
  outputs (chloroplast-membrane.plastoquinol)
  function reduction
  implemented-by psii
  inhibited-by dcmu)
```

Explanation by Pathway Tracing

```
(photosynthesis isa process with
  inputs (chloroplast-inside.water everywhere.light chloroplast-outside.nadph+
          chloroplast-outside.adp chloroplast-outside.pi)
  outputs (chloroplast-outside.atp chloroplast-outside.nadph everywhere.o2)
  implemented-by photosystem)

(photosystem composition (psii antenna-array atpase pq-pool))

(light-absorption isa process with
  inputs (everywhere.light)
  outputs (chlorophyll.energy)
  function absorption
  implemented-by chlorophyll)

(light-energy-concentration isa process with
  outputs psii.energy
  driver chlorophyll.energy
  function concentration
  implemented-by antenna-array)

(psii-water-breakdown isa process with
  inputs (chloroplast-inside.water)
  driver psii.energy
  outputs (psii.e- psii.e- chloroplast-inside.h+ chloroplast-inside.o2)
  function molecular-splitting
  implemented-by psii)

(psii-pq-reduction isa process with
  inputs (psii.e- chloroplast-membrane.h+ chloroplast-membrane.plastoquinone)
  outputs (chloroplast-membrane.plastoquinol)
  function reduction
  implemented-by psii
  inhibited-by dcmu)
```

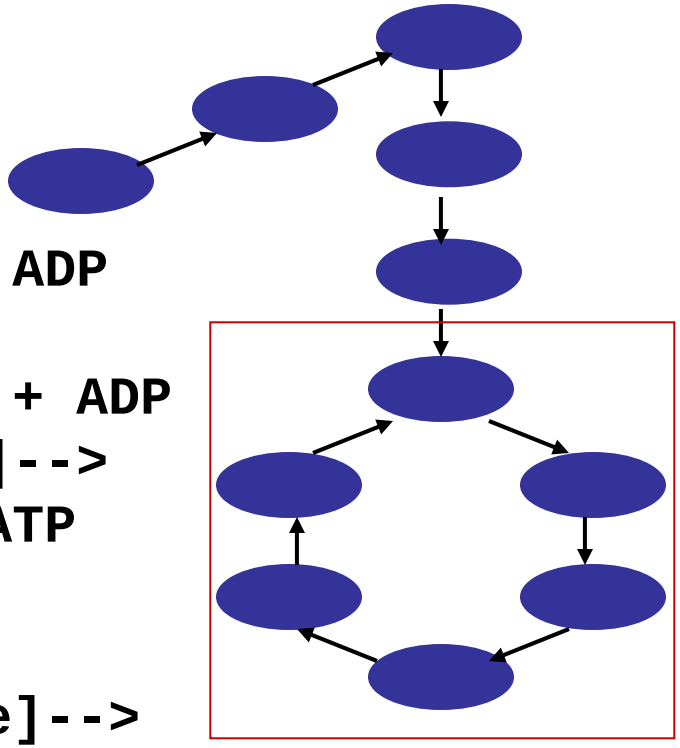


Explanation by Pathway Tracing

```
(track-object 'chloroplast-inside.water)
Tracking CHLOROPLAST-INSIDE.WATER
  -> PHOTOSYNTHESIS:
    Tracking CHLOROPLAST-OUTSIDE.ATP
    Tracking CHLOROPLAST-OUTSIDE.NADPH
    Tracking EVERYWHERE.O2
  -> PSII-WATER-BREAKDOWN:
    Tracking PSII.E-
      -> PSII-PQ-REDUCTION:
        Tracking CHLOROPLAST-MEMBRANE.PLASTOQUINOL
      -> E-FUNNLING-PSII-TO-PSI:
        Tracking PSI.E-
          -> PSI-NADPH-FORMATION:
    Tracking CHLOROPLAST-INSIDE.H+
      -> ATP-FORMATION:
    Tracking CHLOROPLAST-INSIDE.O2
      -> O2-DIFFUSSION:
```

Simulation

Reactions from Glycolysis and the TCA Cycle:



CYTOSOLIC:glucose + ATP
---[Hexokinase]-->
glucose 6-phosphate + ADP

CYTOSOLIC:1,3-bisphosphoglycerate + ADP
---[Phosphoglycerate kinase]-->
3-phosphoglycerate + ATP

MITOCHONDRIAL:isocitrate + NAD+
---[Isocitrate dehydrogenase]-->
a-ketoglutarate + NADH + H+ + Co2

MITOCHONDRIAL:succinyl CoA + GDP + phosphatate
---[Succinyl CoA synthase]-->
succinate + GTP + CoA

Simulation: Find pathways that connect species

Solution for Fructose environment (Target = Malate)

```
frucose ---[Fructokinase]--> fructose 1-phosphate
fructose 1-phosphate ---[Fructose 1-phosphate aldolase]--> glyceraldehyde + dihydrozyacetone phosphate
dihydrozyacetone phosphate ---[Isomerase]--> glyceraldehyde 3-phosphate
phosphatate + NAD+ + glyceraldehyde 3-phosphate ---[Triose phosphate dehydrogenase]--> 1,3-bisphosphoglycerate
1,3-bisphosphoglycerate + ADP ---[Phosphoglycerate kinase]--> 3-phosphoglycerate + ATP
3-phosphoglycerate ---[Phosphoglyceromutase]--> 2-phosphoglycerate
2-phosphoglycerate ---[Enolase]--> phosphoenolpyruvate + H2O
phosphoenolpyruvate + ATP ---[Pyruvate kinase]--> pyruvate + ADP
malate + NAD+ ---[Malate dehydrogenase]--> oxaloacetate + NADH + H+
pyruvate + NAD+ + CoA ---[NIL]--> NADH + H+ + Co2 + acetyl CoA
acetyl CoA + oxaloacetate ---[Citrate synthase]--> citrate + CoA
citrate ---[Aconitase]--> isocitrate
isocitrate + NAD+ ---[Isocitrate dehydrogenase]--> a-ketoglutarate + NADH + H+ + Co2
a-ketoglutarate + NAD+ + CoA ---[a-ketogluterate dehydrogenase complex]--> succinyl CoA + NADH + H+ + Co2
succinyl CoA + GDP + phosphatate ---[Succinyl CoA synthase]--> succinate + GTP + CoA
succinate + FAD ---[Succinate dehydrogenase]--> fumarate + FADH2
fumarate + H2O ---[Fumerase]--> malate
```

Solution for Glucose environment (Target = Malate)

```
glucose + ATP ---[Hexokinase]--> glucose 6-phosphate + ADP
glucose 6-phosphate ---[Phosphoglucomutase]--> frucose 6-phosphate
frucose 6-phosphate + ATP ---[Phosphofructokinase]--> frucose 1,6 bisphosphate + ADP
frucose 1,6 bisphosphate ---[Aldolase]--> dihydrozyacetone phosphate + glyceraldehyde 3-phosphate
phosphatate + NAD+ + glyceraldehyde 3-phosphate ---[Triose phosphate dehydrogenase]--> 1,3-bisphosphoglycerate
1,3-bisphosphoglycerate + ADP ---[Phosphoglycerate kinase]--> 3-phosphoglycerate + ATP
…[same as above from this point onward…]
```

Simulation: Simulate natural or experimental “knockouts”...

```
glucose + ATP ---[Hexokinase]--> glucose 6-phosphate + ADP
glucose 6-phosphate ---[Phosphoglucomutase]--> frucose 6-phosphate
frucose 6-phosphate + ATP ---[Phosphofructokinase]--> frucose 1,6 bisphosphate + ADP
frucose 1,6 bisphosphate ---[Aldolase]--> dihydrozyacetone phosphate + glyceraldehyde 3-phosphate
phosphatate + NAD+ + glyceraldehyde 3-phosphate ---[Triose phosphate dehydrogenase]--> 1,3-bisphosphoglycerate
1,3-bisphosphoglycerate + ADP ---[Phosphoglycerate kinase]--> 3-phosphoglycerate + ATP
3-phosphoglycerate ---[Phosphoglyceromutase]--> 2-phosphoglycerate
2-phosphoglycerate ---[Enolase]--> phosphoenolpyruvate + H2O
phosphoenolpyruvate + ATP ---[Pyruvate kinase]--> pyruvate + ADP
malate + NAD+ ---[Malate dehydrogenase]--> oxaloacetate + NADH + H+
pyruvate + NAD+ + CoA ---[NIL]--> NADH + H+ + Co2 + acetyl CoA
acetyl CoA + oxaloacetate ---[Citrate synthase]--> citrate + CoA
citrate ---[Aconitase]--> isocitrate
isocitrate + NAD+ ---[Isocitrate dehydrogenase]--> a-ketoglutarate + NADH + H+ + Co2
a-ketoglutarate + NAD+ + CoA ---[a-ketogluterate dehydrogenase complex]--> succinyl CoA + NADH + H+ + Co2
succinyl CoA + GDP + phosphatate ---[Succinyl CoA synthase]--> succinate + GTP + CoA
succinate + FAD ---[Succinate dehydrogenase]--> fumarate + FADH2
fumarate + H2O ---[Fumerase]--> malate
```

Knockout:

```
1,3-bisphosphoglycerate + ADP
      ---[Phosphoglycerate kinase]-->
                          3-phosphoglycerate + ATP
```

Simulation: ...and propose “bridging” reactions

Knockout:

```
1,3-bisphosphoglycerate + ADP
      ---[Phosphoglycerate kinase]-->
                          3-phosphoglycerate + ATP
```

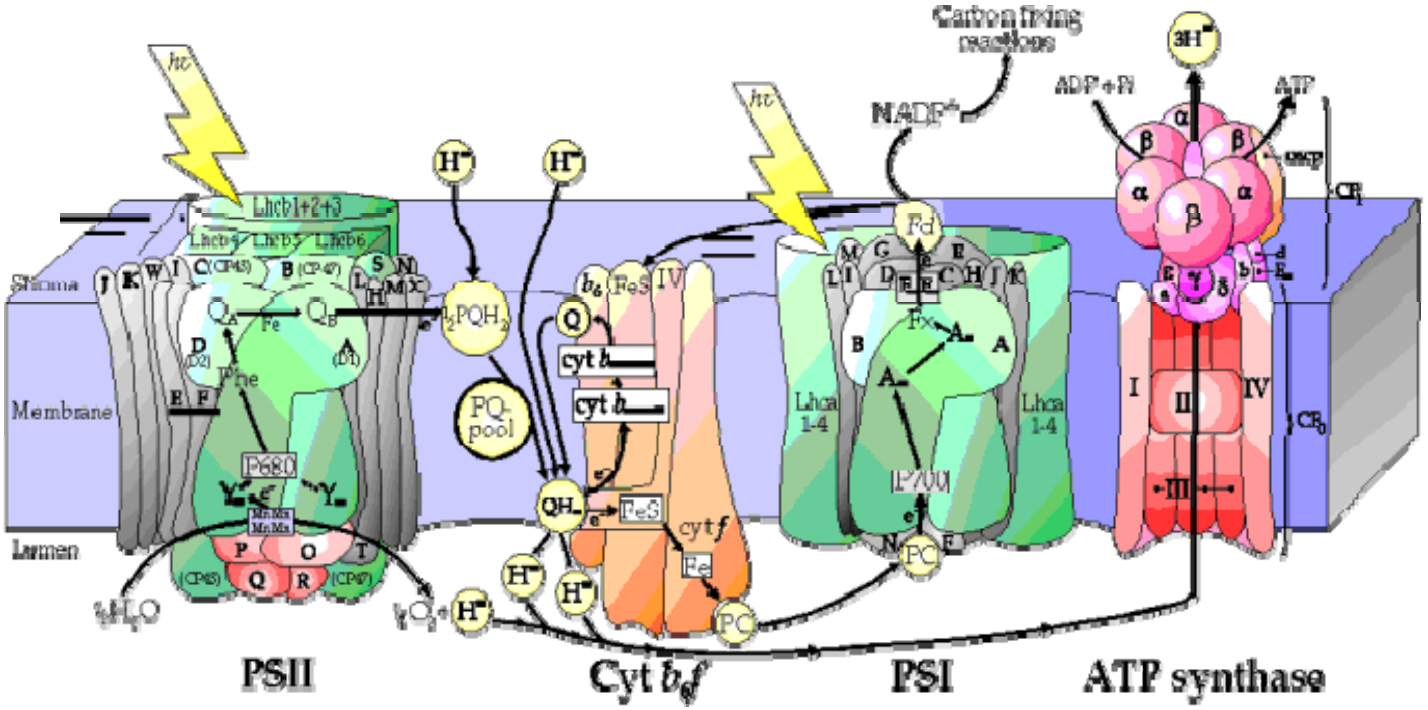
25 plausible (single) “bridging” reactions are proposed:

```
<CYTOSOLIC:glyceraldehyde 3-phosphate ---[]--> 3-phosphoglycerate>
<CYTOSOLIC:dihydrozyacetone phosphate ---[]--> 3-phosphoglycerate>
<CYTOSOLIC:frucose 1,6 bisphosphate ---[]--> phosphoenolpyruvate + 3-phosphoglycerate>
<CYTOSOLIC:frucose 1,6 bisphosphate ---[]--> 2-phosphoglycerate + 3-phosphoglycerate>
<CYTOSOLIC:frucose 1,6 bisphosphate ---[]--> 3-phosphoglycerate + 3-phosphoglycerate>
<CYTOSOLIC:ATP + frucose 1,6 bisphosphate ---[]--> ADP + 1,3-bisphosphoglycerate + 3-phosphoglycerate>
<CYTOSOLIC:frucose 1,6 bisphosphate ---[]--> glyceraldehyde 3-phosphate + 3-phosphoglycerate>
<CYTOSOLIC:frucose 1,6 bisphosphate ---[]--> dihydrozyacetone phosphate + 3-phosphoglycerate>
<CYTOSOLIC:ADP + frucose 1,6 bisphosphate ---[]--> ATP + Co2 + acetyl + 3-phosphoglycerate>
<CYTOSOLIC:ADP + 1,3-bisphosphoglycerate
                  ---[]--> ATP + 3-phosphoglycerate>
<CYTOSOLIC:ADP + frucose 1,6 bisphosphate ---[]--> ATP + pyruvate + 3-phosphoglycerate>
<CYTOSOLIC:ADP + frucose 1,6 bisphosphate ---[]--> ATP + glycerate + 3-phosphoglycerate>
<CYTOSOLIC:ADP + frucose 1,6 bisphosphate ---[]--> ATP + glyceraldehyde + 3-phosphoglycerate>
<CYTOSOLIC:ADP + frucose 1,6 bisphosphate ---[]--> ATP + dihydroxyacetone + 3-phosphoglycerate>
<CYTOSOLIC:ATP + glucose 6-phosphate ---[]--> ADP + phosphoenolpyruvate + 3-phosphoglycerate>
<CYTOSOLIC:ATP + glucose 6-phosphate ---[]--> ADP + 2-phosphoglycerate + 3-phosphoglycerate>
<CYTOSOLIC:ATP + glucose 6-phosphate ---[]--> ADP + 3-phosphoglycerate + 3-phosphoglycerate>
<CYTOSOLIC:ATP + glucose 6-phosphate ---[]--> ADP + glyceraldehyde 3-phosphate + 3-phosphoglycerate>
<CYTOSOLIC:ATP + glucose 6-phosphate ---[]--> ADP + dihydrozyacetone phosphate + 3-phosphoglycerate>
<CYTOSOLIC:glucose 6-phosphate ---[]--> Co2 + acetyl + 3-phosphoglycerate>
<CYTOSOLIC:glucose 6-phosphate ---[]--> pyruvate + 3-phosphoglycerate>
<CYTOSOLIC:glucose 6-phosphate ---[]--> glycerate + 3-phosphoglycerate>
<CYTOSOLIC:glucose 6-phosphate ---[]--> glyceraldehyde + 3-phosphoglycerate>
<CYTOSOLIC:glucose 6-phosphate ---[]--> dihydroxyacetone + 3-phosphoglycerate>
<CYTOSOLIC:glucose + ATP ---[]--> 1,3-bisphosphoglycerate + 3-phosphoglycerate>
```

Simulation: What does this model predict?

Explanation: How does it make these predictions?

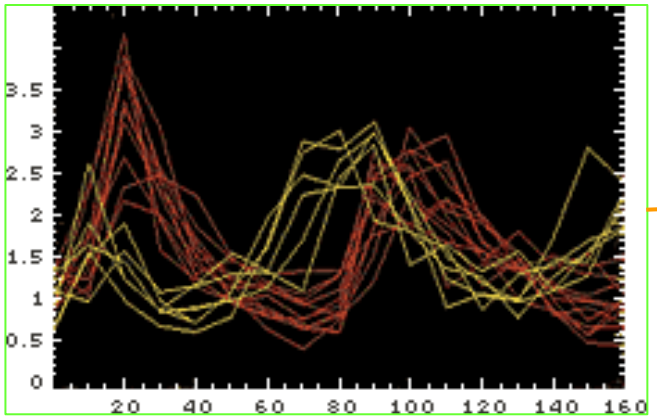
Model Identification: What models fit this data?



Model formation and revision

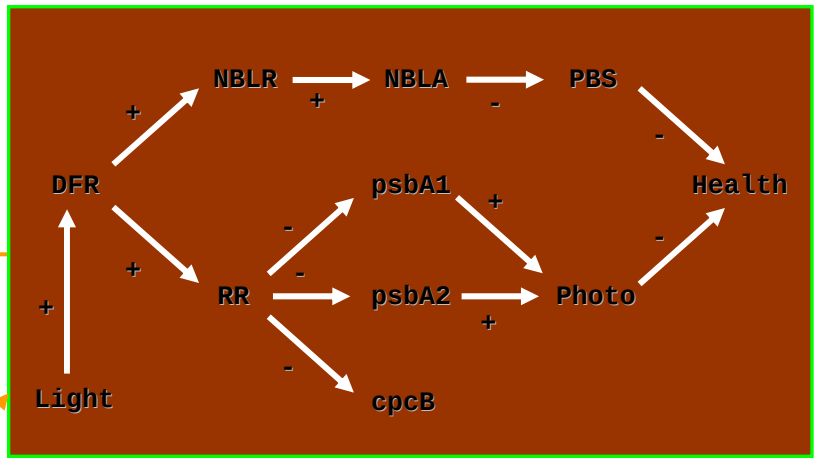
Interactive Guidance from Scientists

Experimental data

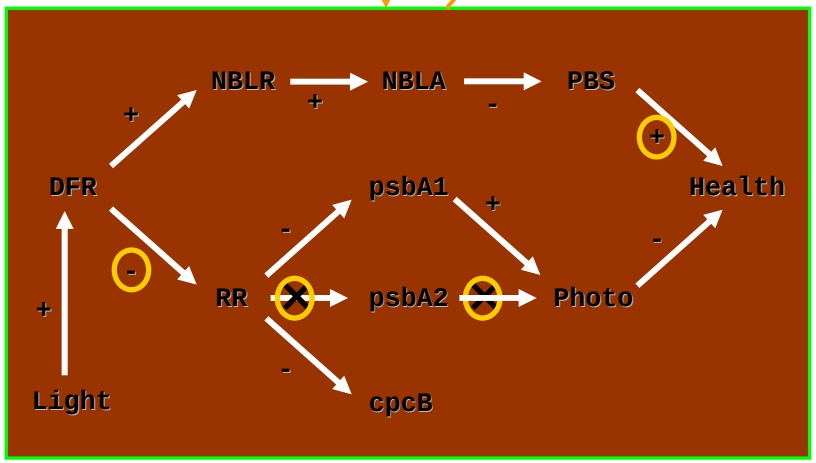


Discovery

Background knowledge

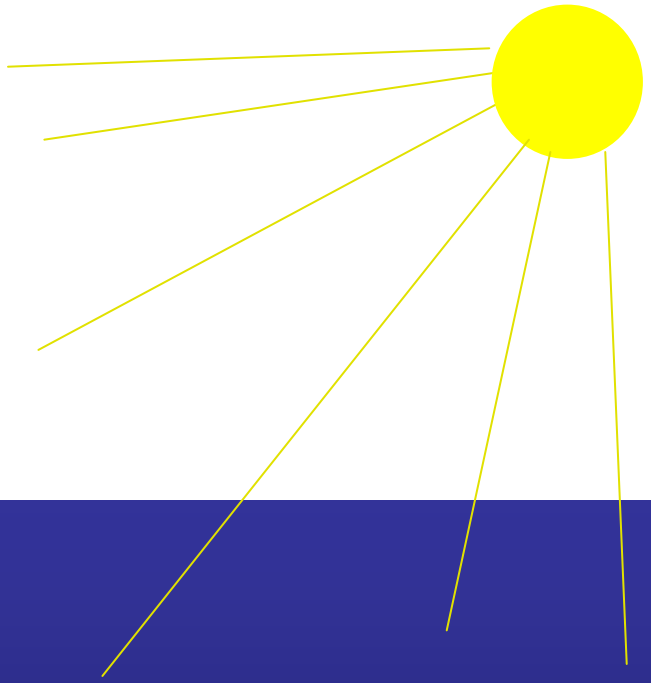


Updated models

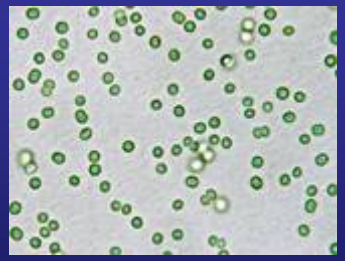


How do cells control response to light?

I.e., What genes are related to the adaptation to high light?

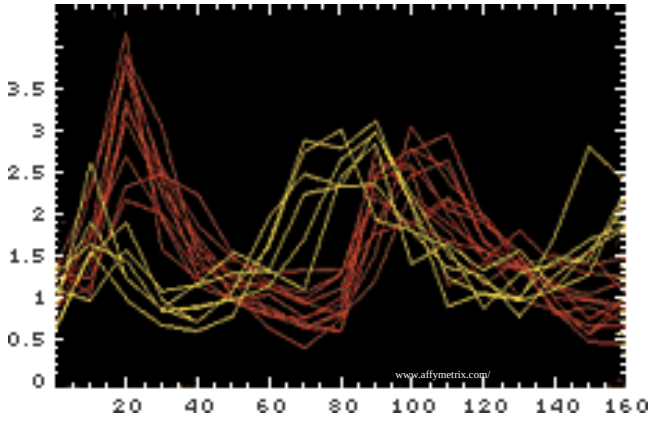


Prochlorococcus MED4

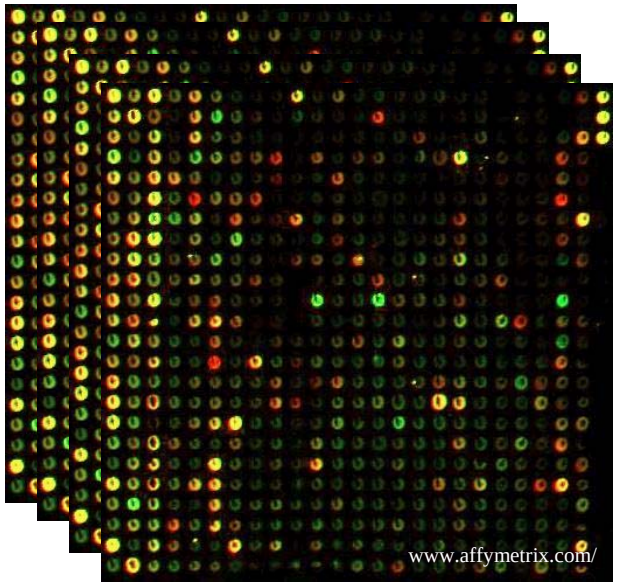


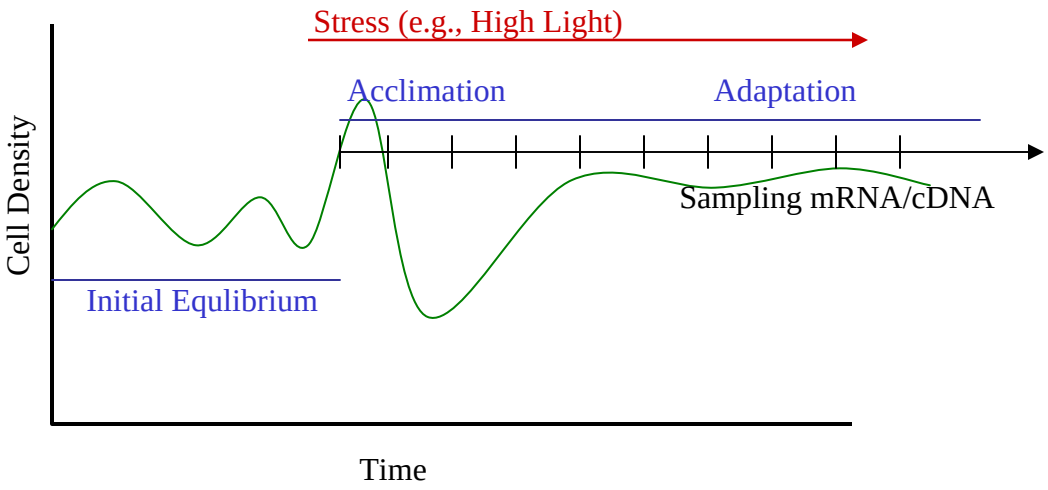
Prochlorococcus MIT9313

The Data: Analyzing Acclimation Dynamics

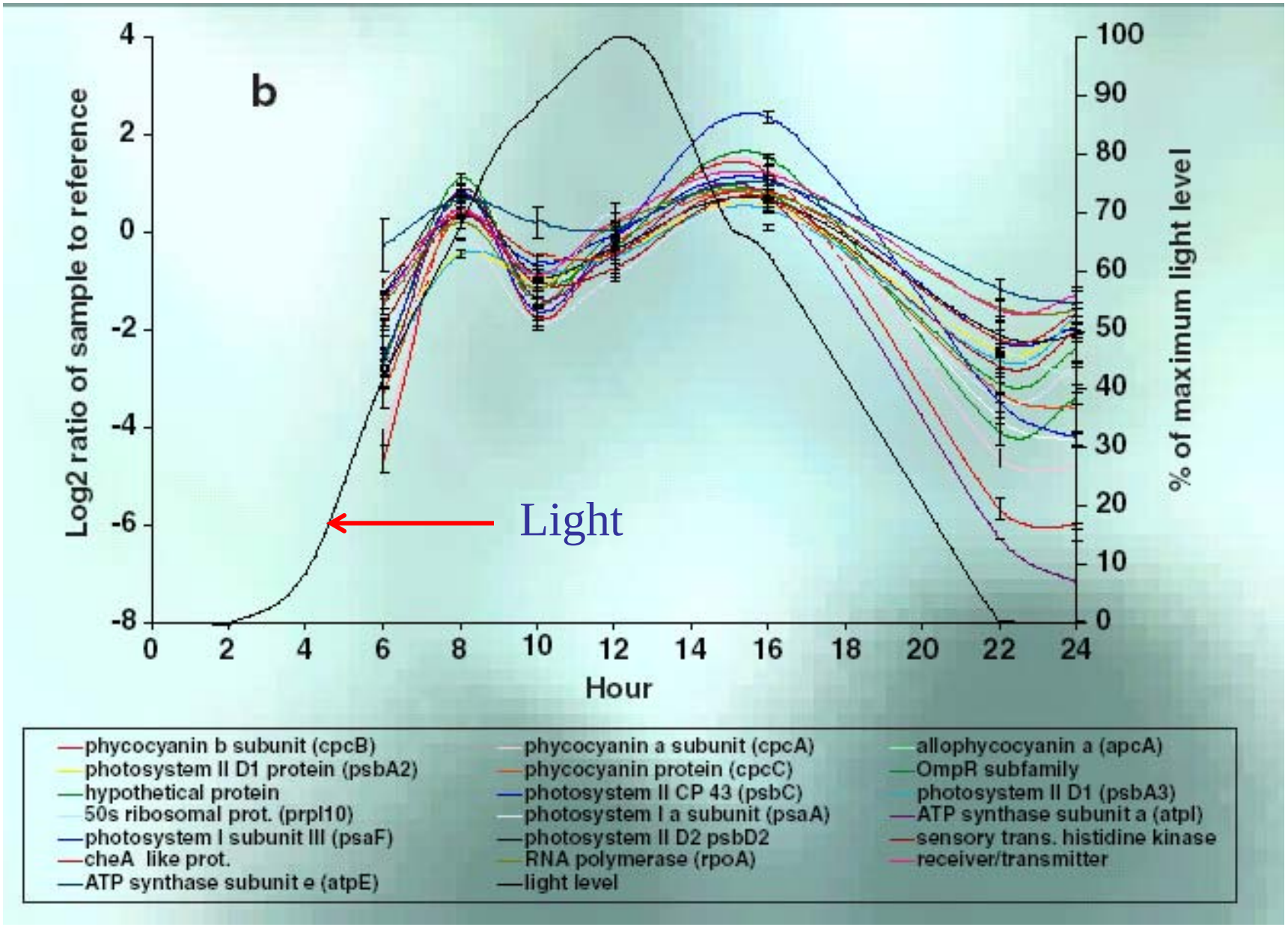


Statistical Annotation

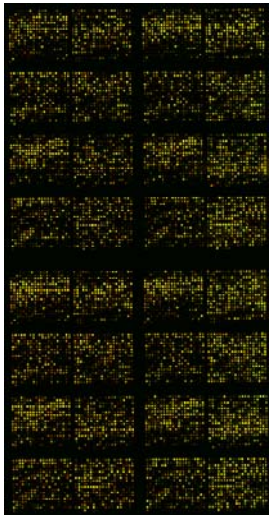




Most positively light-correlated responses:



Model formation and revision

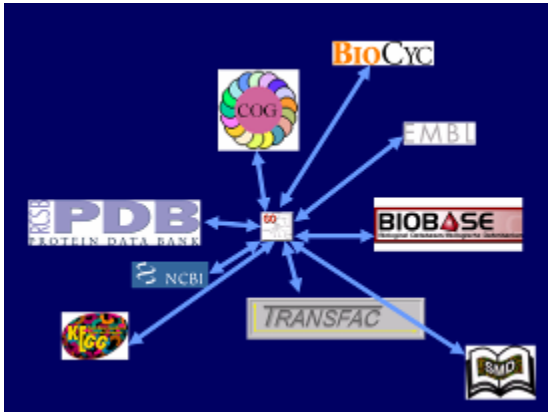


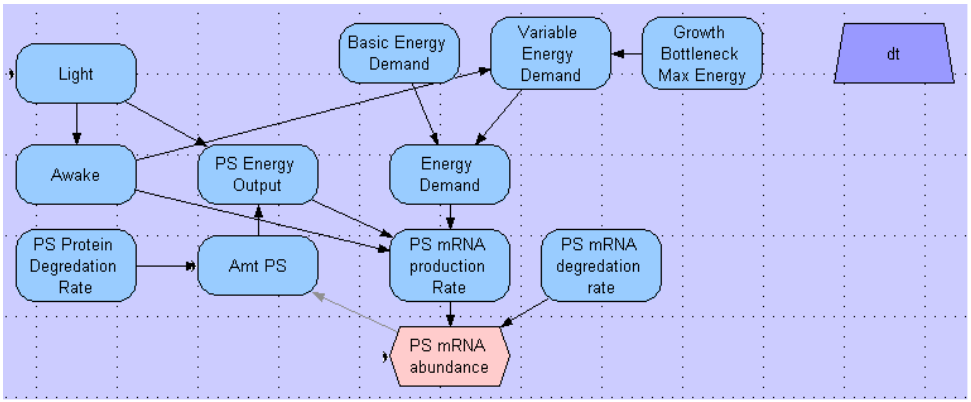
Statistics (R)

Constraints

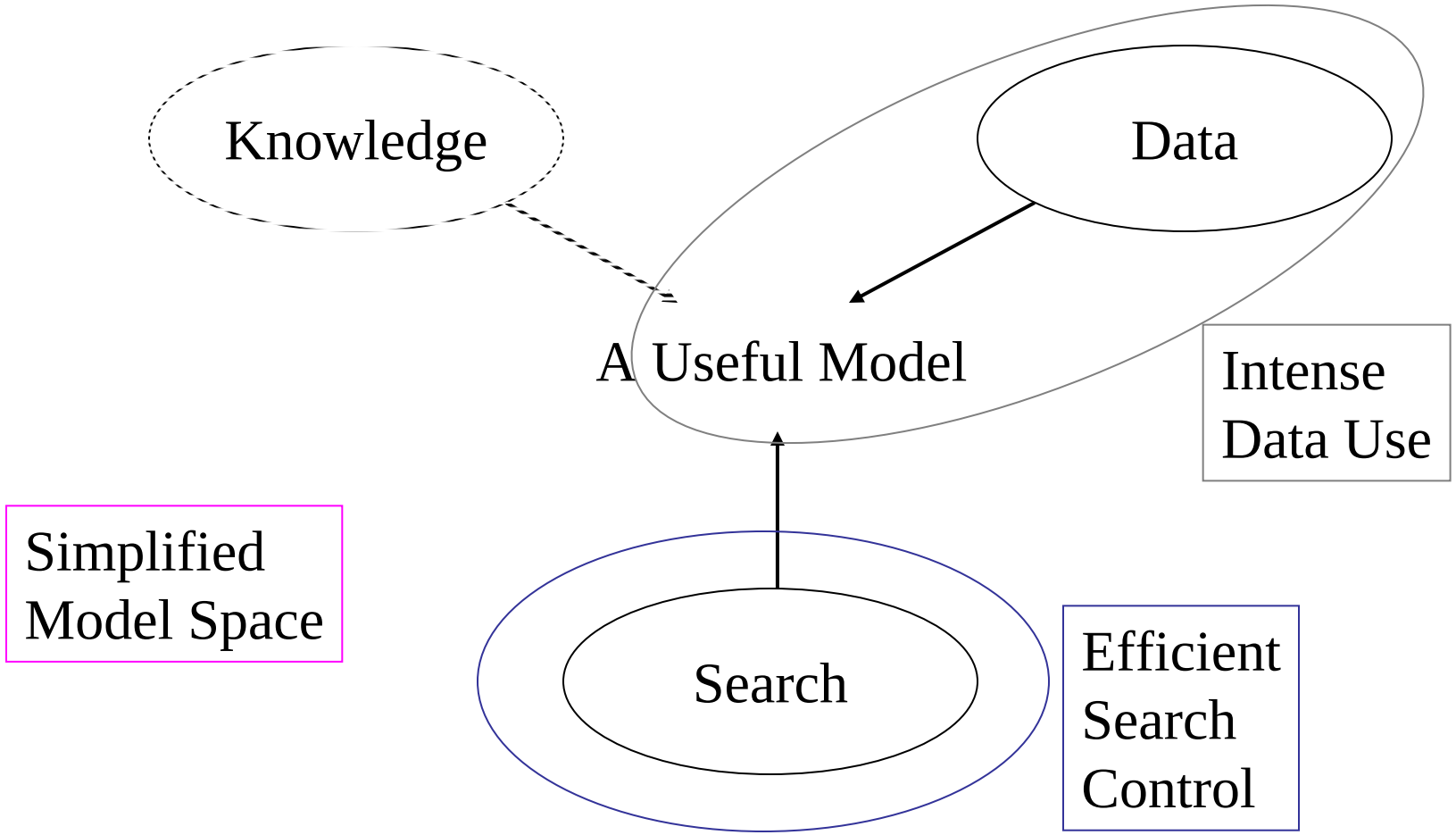
Knowledge Summarization

Model Space Search

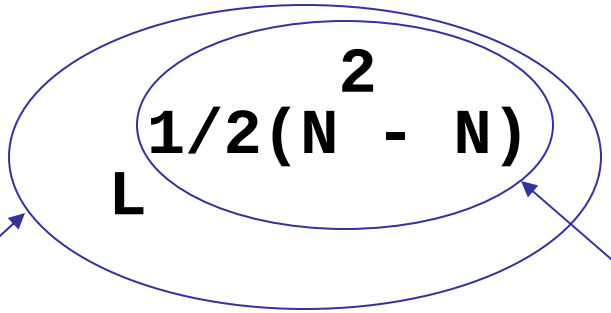




“Knowledge lean” (*de novo*) Discovery



How many regulatory models are there for n genes (In the worst case)?



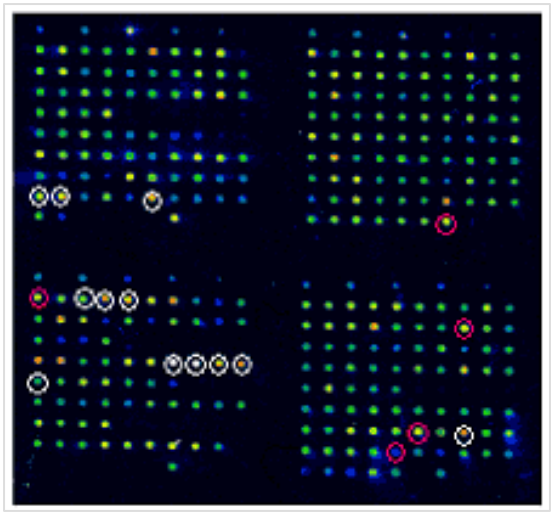
Number of combinations of L link types

Number of ways to arrange links among N nodes

How many regulatory models are there for n genes

(In the worst case)?

$$L^{1/2(N^2 - N)}$$



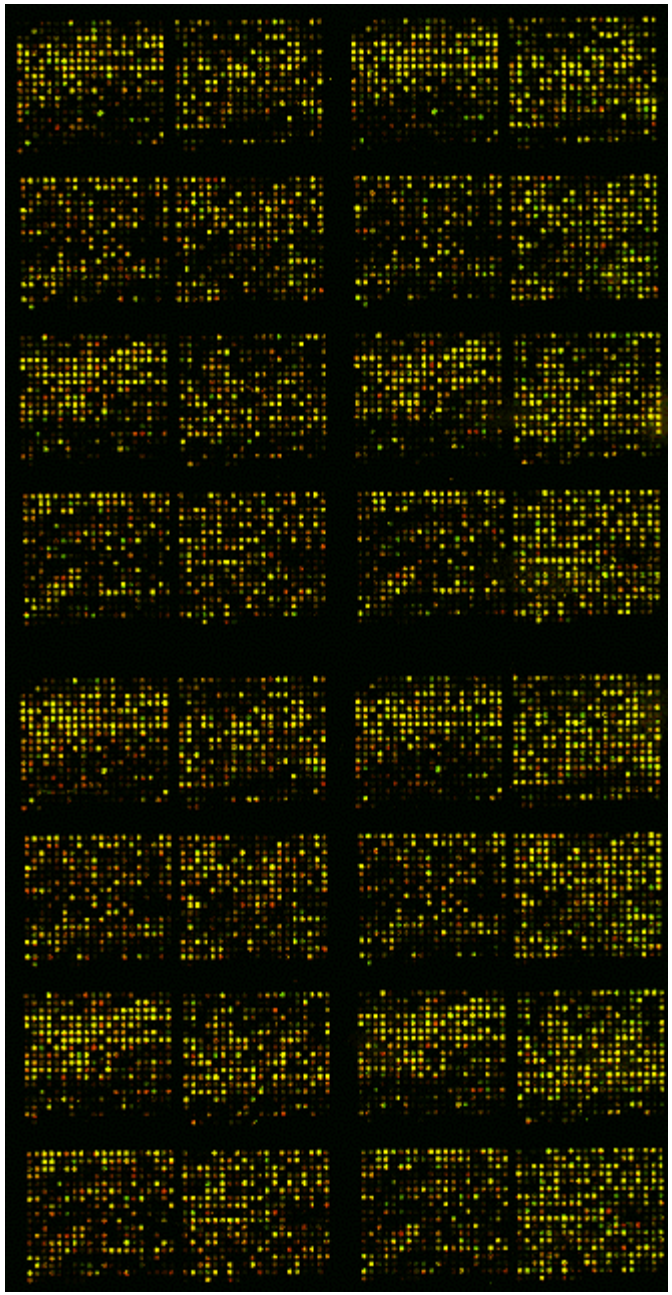
N=300
L=4

$$4^{1/2(300^2 - 300)}$$

$$\sim 4^{89700}$$

$$\sim \textit{Infinity}$$

Identification requires ~2^n observations!



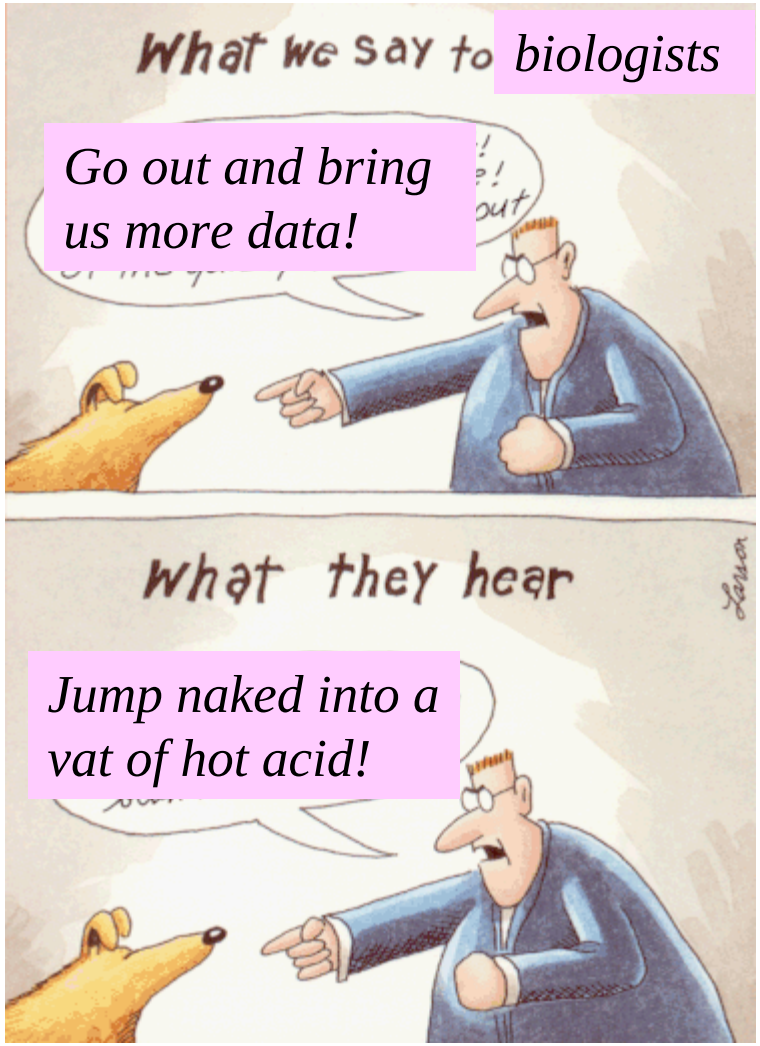
How many models are there
for the C. reinhardtii chip?

N=8000
L=4

$$\sim 4^{1/2(8000^2 - 8000)}$$

$$\sim 4^{31996000}$$

(Not to mention 2^{8000} observations!)

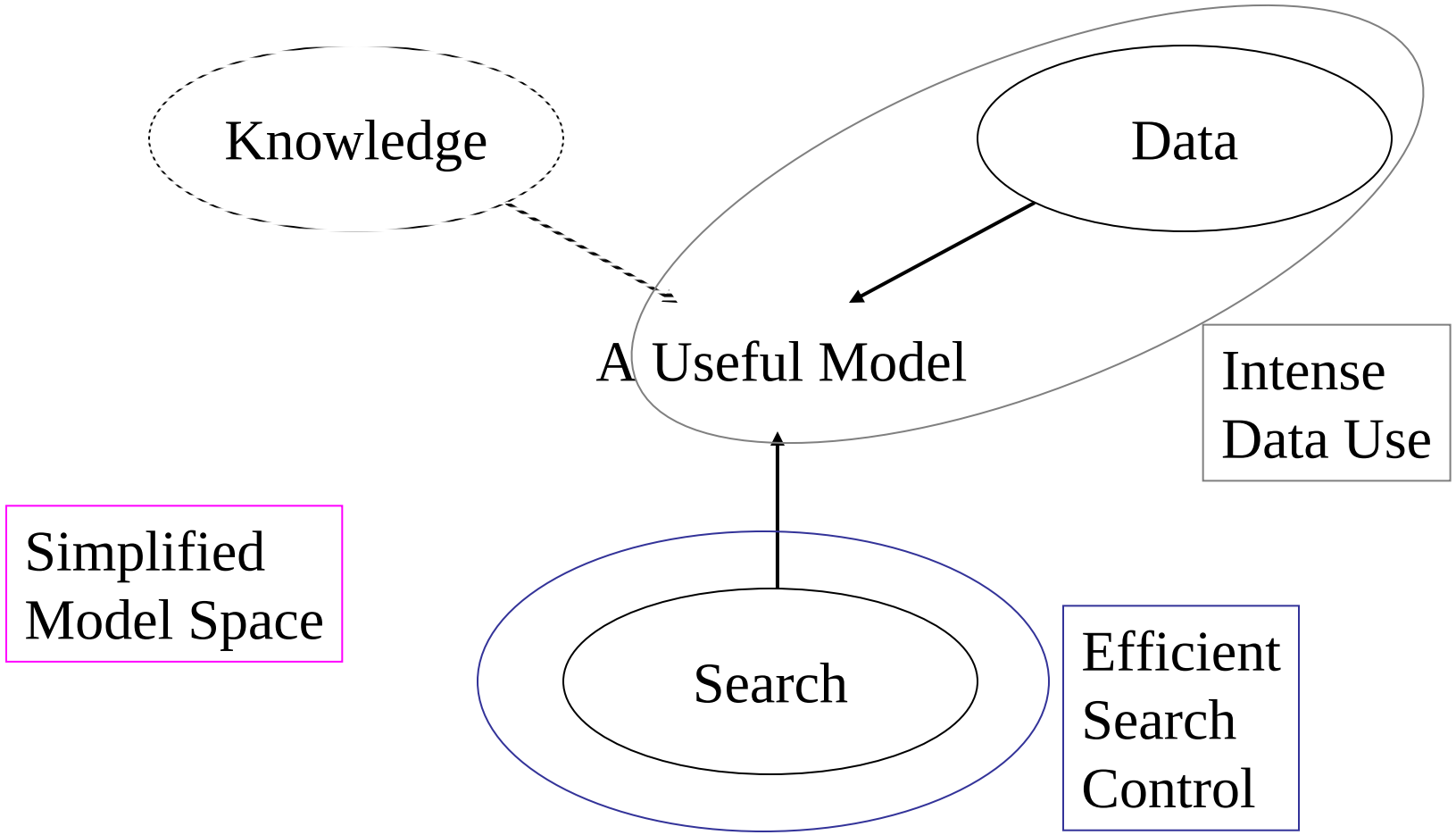

What we say to biologists
Go out and bring us more data!
What they hear
Jump naked into a vat of hot acid!

Shrager's first law of (computational) biology:

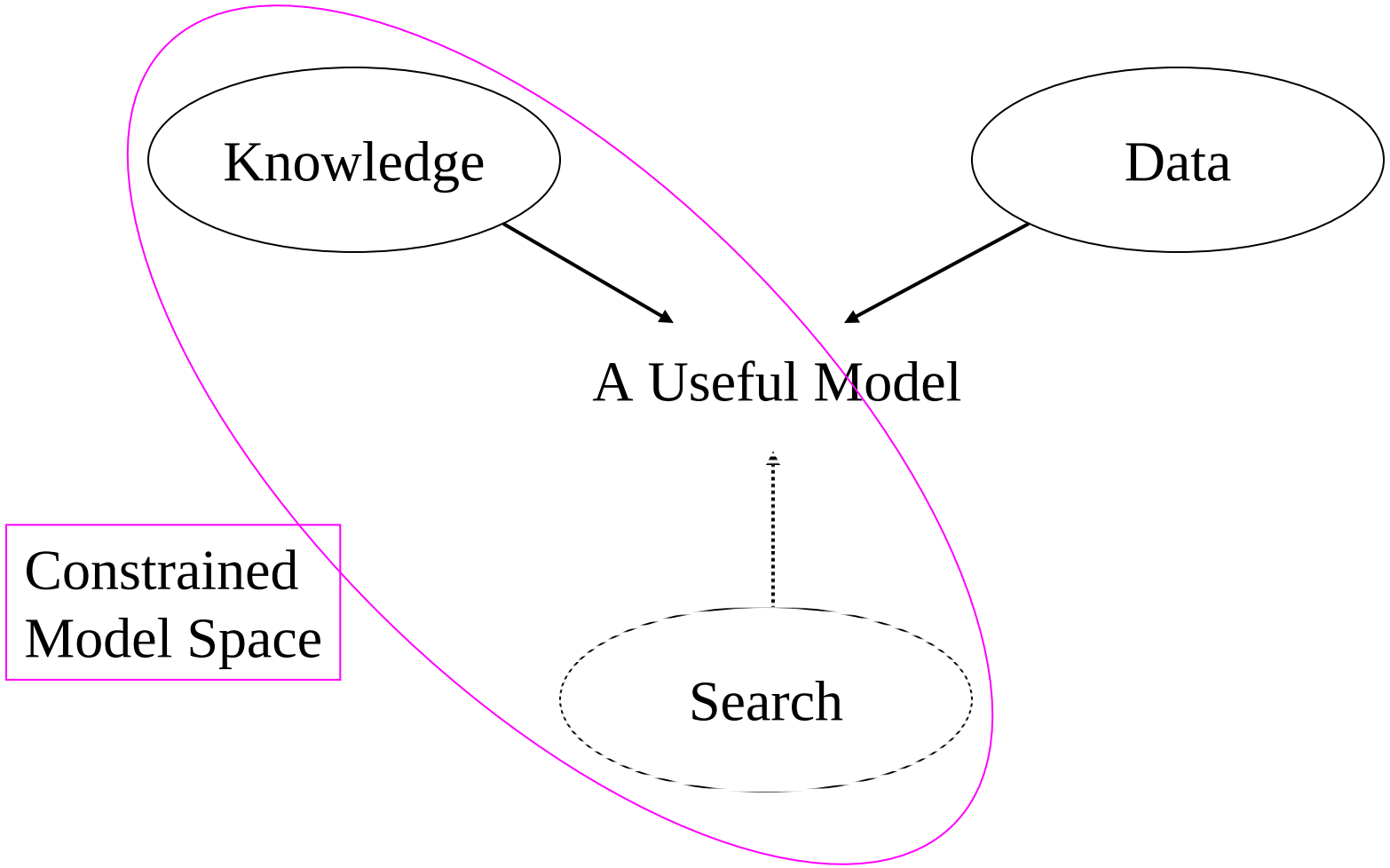
If you think that you need more data…..

You need more knowledge!

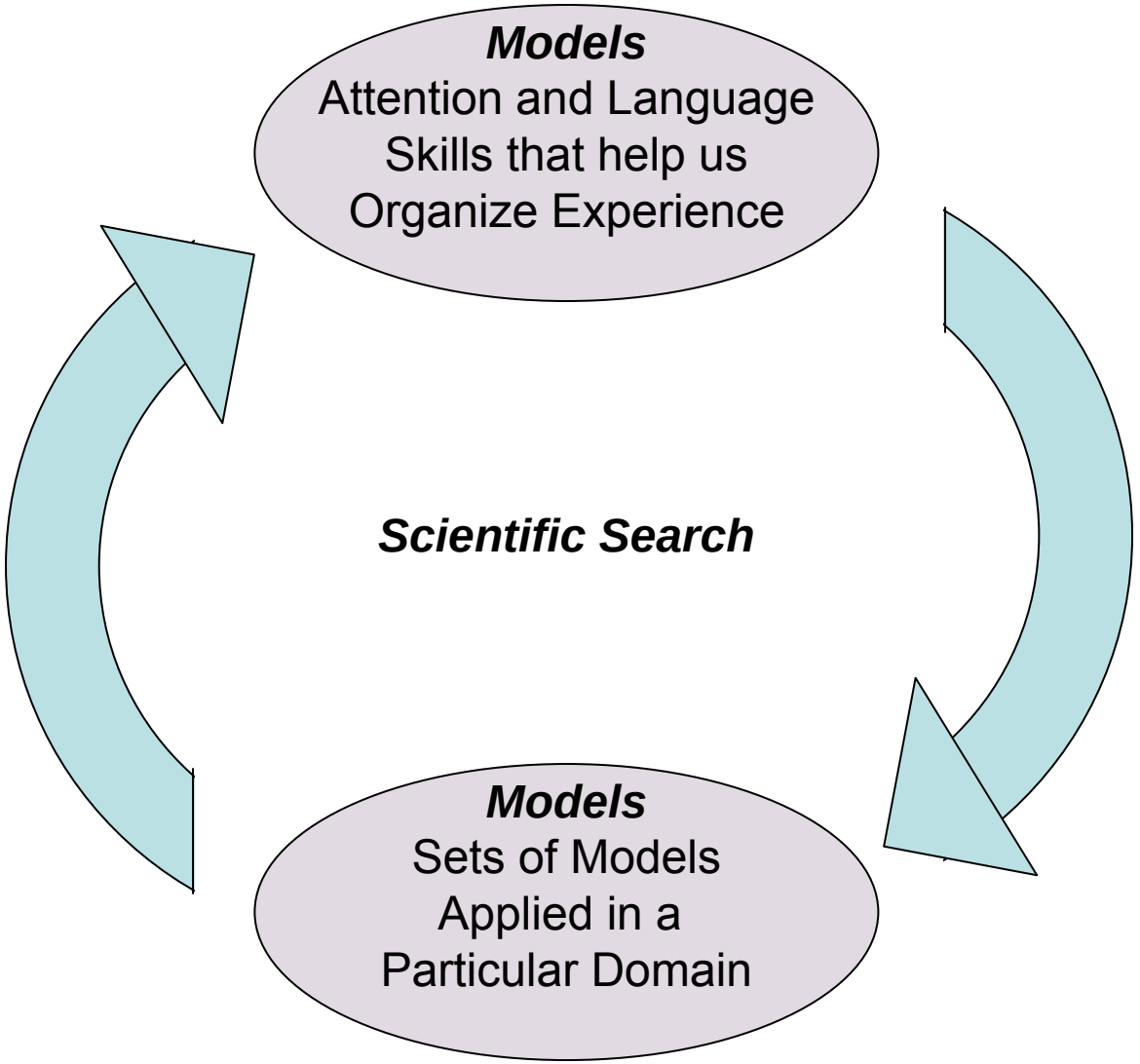
“Knowledge lean” (*de novo*) Discovery

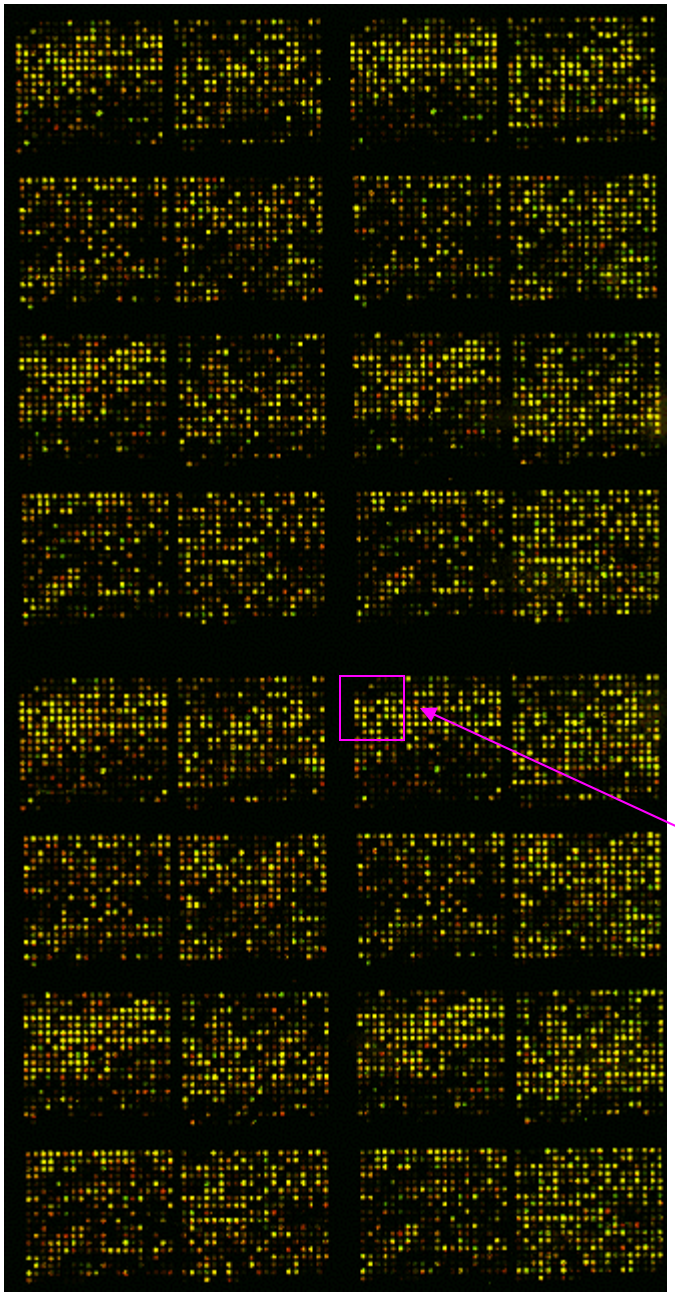


“Knowledge Rich” Computational Discovery

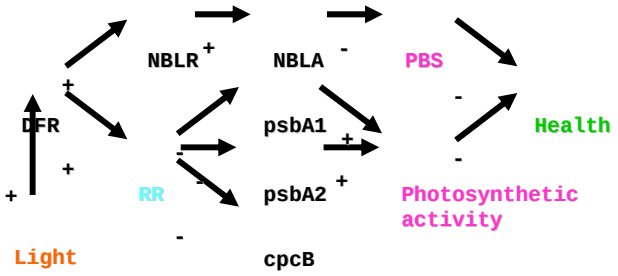


Explanation is the main function of theories (models)





Adding knowledge: Limiting search to subsystems.



Adding Knowledge: Annotate the theory in terms of Models.

What are Models?

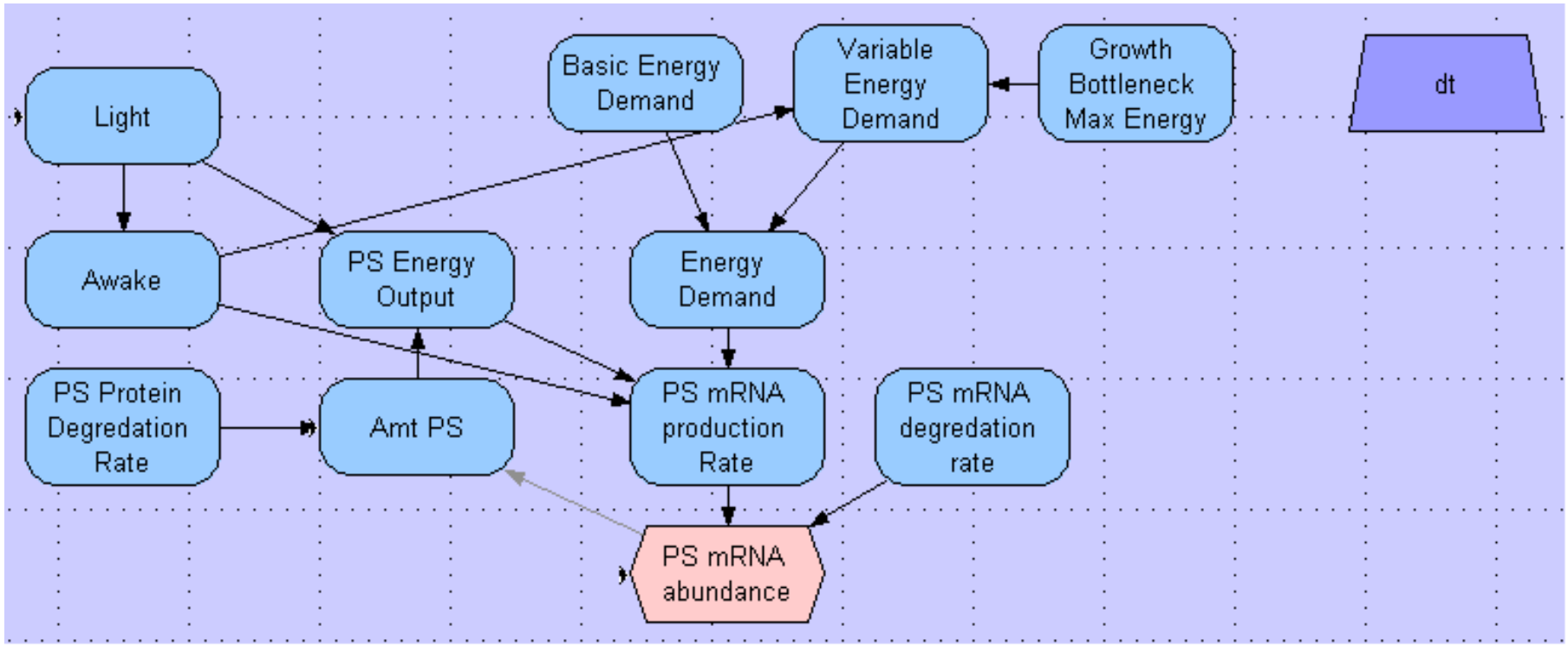
Conceptually coherent, possibly complex, units of partially abstract knowledge that can be incrementally "mixed into" an existing model (by "Model Application"), updating the model in accord with the principles represented in the Model.

(aka. Schemas, Scripts)

Some Models in Cell Biology:

Transcriptional Regulation
Attentuation
Transposon Insertion
Allosteric Modulation
Signal Transduction
Operon
Chemical Cycle
Feedback Regulation
Protein Assembly

Graphical Model for Light Response Curve:



variables $light, mRNA, photo\_protein, ros, redox, trans$
observables $light, mRNA$;

process photosynthesis;
  equations $d[redox, t, 1] = 0.01 * light * photo\_protein$;
  $d[ros, t, 1] = 0.001 * light * photo\_protein$;

process photo_translation;
  equations $d[photo\_protein, t, 1] = 0.9 * mRNA$;

process protein_degradation_ros;
  conditions $photo\_protein > 0$;
  equations $d[photo\_protein, t, 1] = -0.05 * ros$;
  $d[ros, t, 1] = -0.05 * ros$;

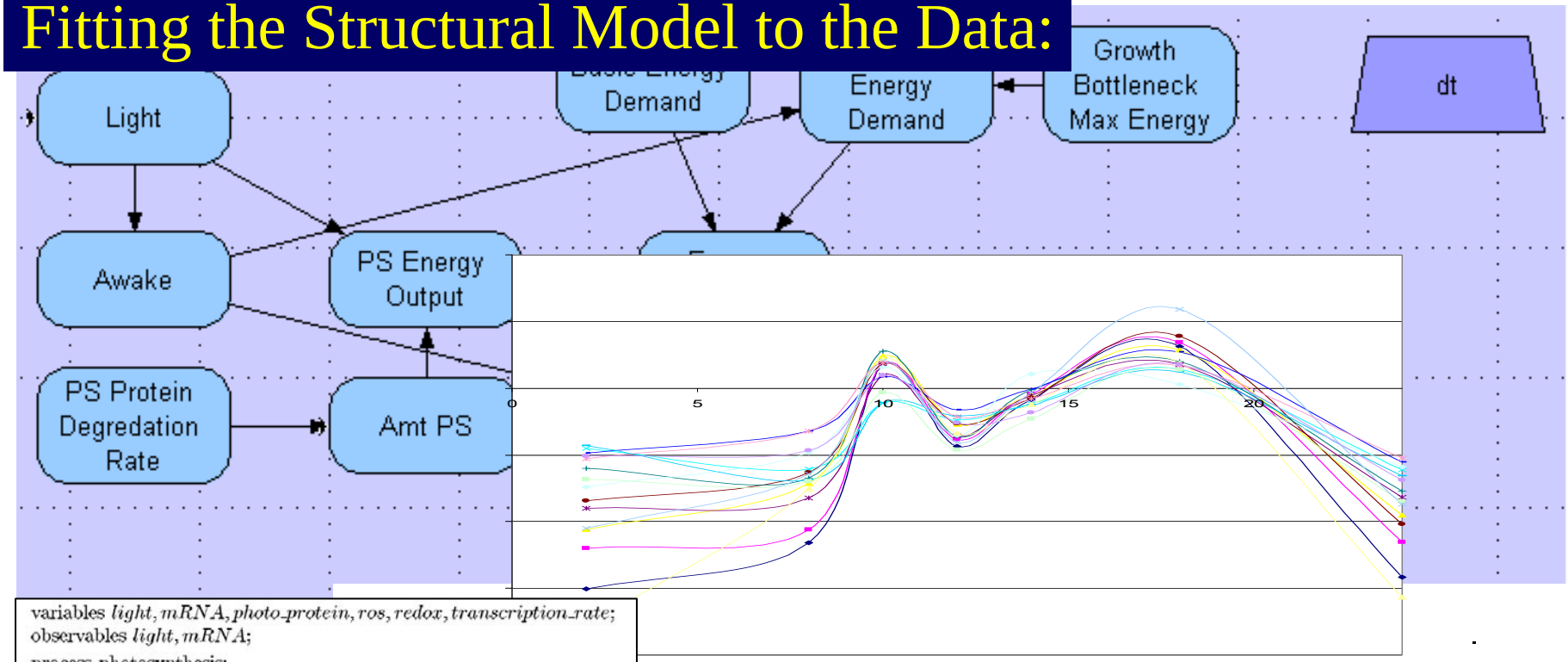
process mRNA_transcription;
  equations $d[mRNA, t, 1] = transcription\_rate$;

process transcription_up;
  equations $transcription\_rate = 0.5 * light$;

process transcription_down;
  conditions $redox > 0$;
  equations $transcription\_rate = -0.1 * redox$;
  $d[redox, t, 1] = -0.01 * redox$;

process mRNA_degradation;
  equations $d[mRNA, t, 1] = -0.01 * mRNA$;

Fitting the Structural Model to the Data:



variables $light, mRNA, photo\_protein, ros, redox, transcription\_rate$;
observables $light, mRNA$;

process photosynthesis;
  equations $d[redox, t, 1] = 0.01 * light * photo\_protein$;
  $d[ros, t, 1] = 0.001 * light * photo\_protein$;

process photo_translation;
  equations $d[photo\_protein, t, 1] = 0.9 * mRNA$;

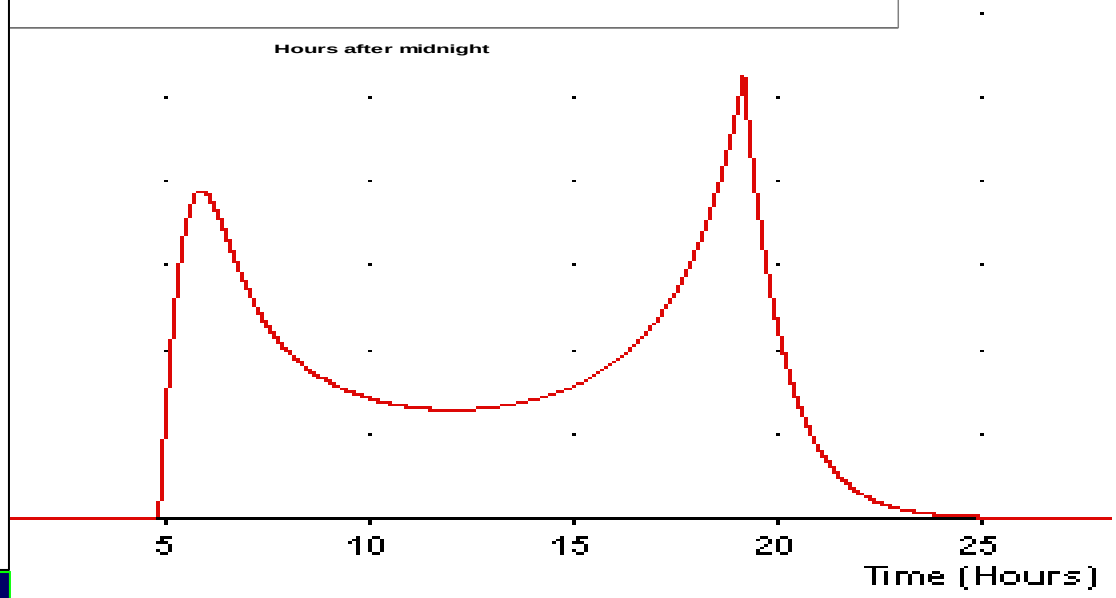
process protein_degradation_ros;
  conditions $photo\_protein > 0$;
  equations $d[photo\_protein, t, 1] = -0.05 * ros$;
  $d[ros, t, 1] = -0.05 * ros$;

process mRNA_transcription;
  equations $d[mRNA, t, 1] = transcription\_rate$;

process transcription_up;
  equations $transcription\_rate = 0.5 * light$;

process transcription_down;
  conditions $redox > 0$;
  equations $transcription\_rate = -0.1 * redox$;
  $d[redox, t, 1] = -0.01 * redox$;

process mRNA_degradation;
  equations $d[mRNA, t, 1] = -0.01 * mRNA$;



Unparameterized (Unfitted) Model

Parameterized (*Fitted*) Model

model Photo_Reg;

variables $light, mRNA, photo\_protein, ROS, redox, transcriptio$
observables $light, mRNA$;
initials $mRNA = 0.253, photo\_protein = 0.836, ROS = 0.059, re$

process photosynthesis;
  equations $d[redox, t, 1] = 0.0155 * light * photo\_protein$;
            $d[ros, t, 1] = 0.019 * light * photo\_protein$;

process photo_translation;
  equations $d[photo\_protein, t, 1] = 7.539 * mRNA$;

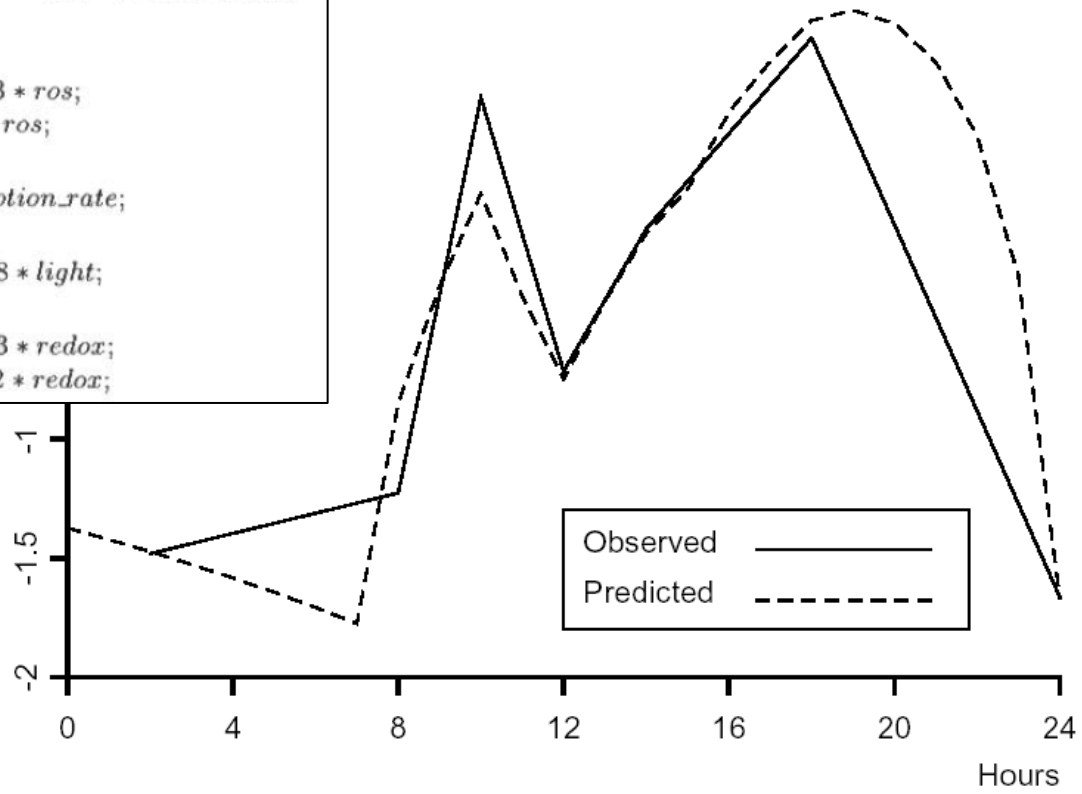
process automatic_degradation1;
  conditions $photo\_protein > 0$;
  equations $d[photo\_protein, t, 1] = -1 * 1.905 * photo\_protein$;

process controlled_degradation1;
  conditions $redox > 0, ros > 0$;
  equations $d[redox, t, 1] = -1 * 0.0003 * ros$;
            $d[ros, t, 1] = -1 * 0.0003 * ros$;

process mRNA_transcription;
  equations $d[mRNA, t, 1] = transcription\_rate$;

process regulate_one_1;
  equations $transcription\_rate = 0.938 * light$;

process regulate_two_2;
  equations $transcription\_rate = 1.203 * redox$;
            $d[redox, t, 1] = -1 * 0.0002 * redox$;

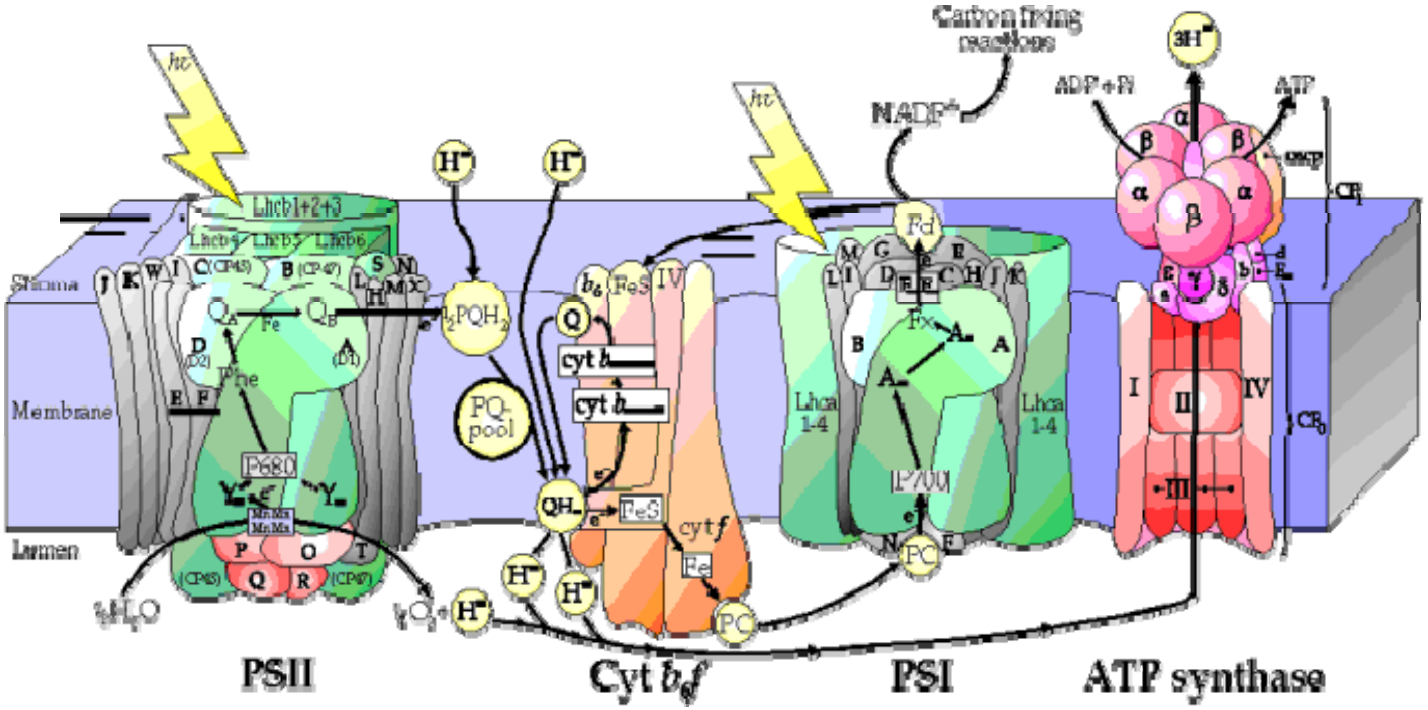


Predicted and observed levels of average gene expression over a 24-hour period.

Simulation: What does this model predict?

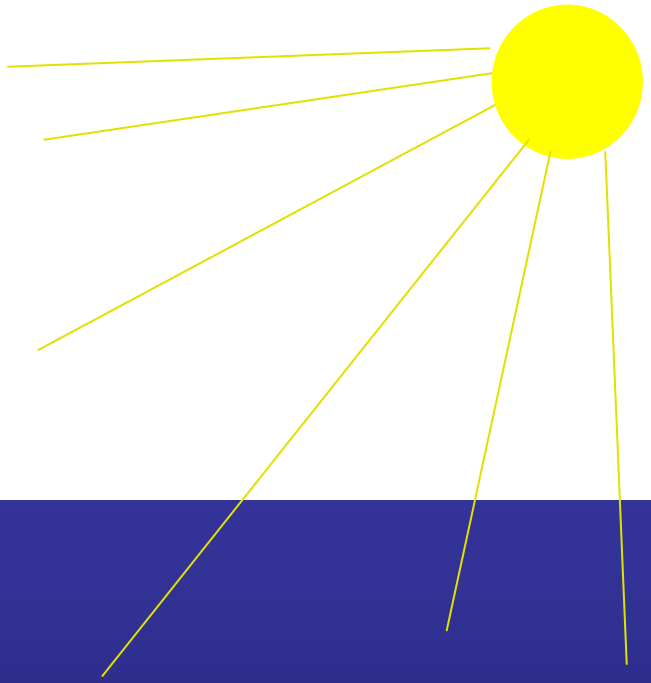
Explanation: How does it make these predictions?

Model Identification: What models fit this data?

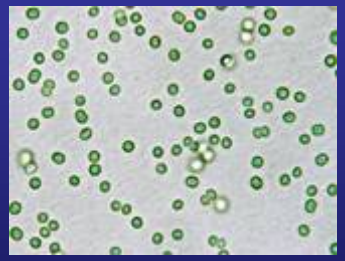


How do cells control response to light?

I.e., What genes are related to the adaptation to high light?

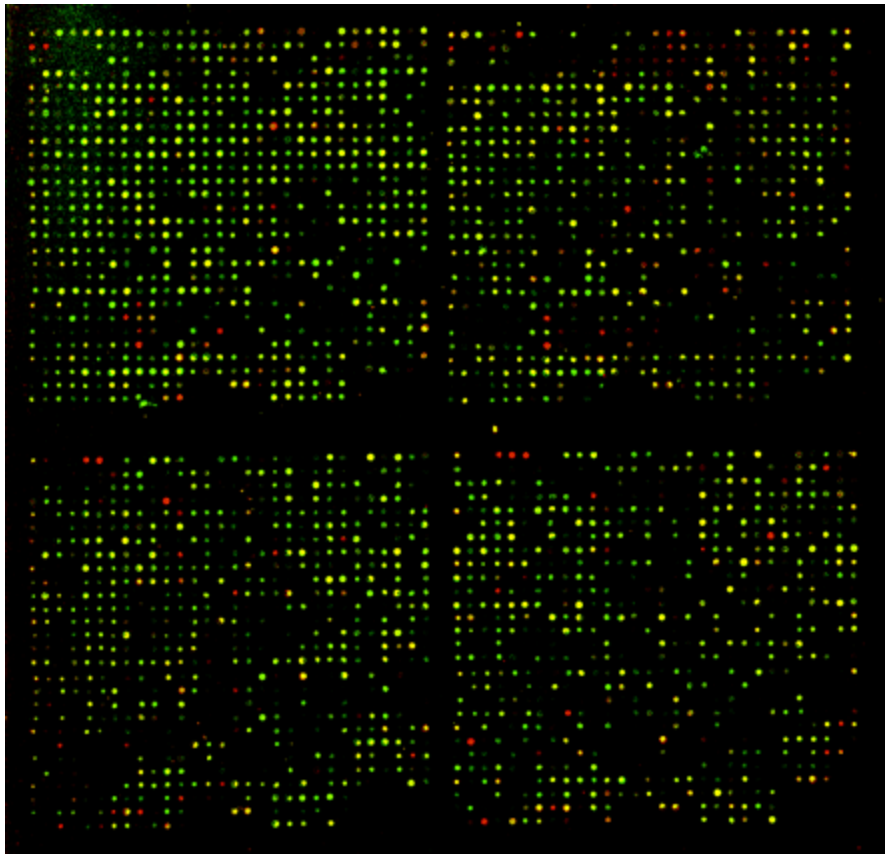


Prochlorococcus MED4

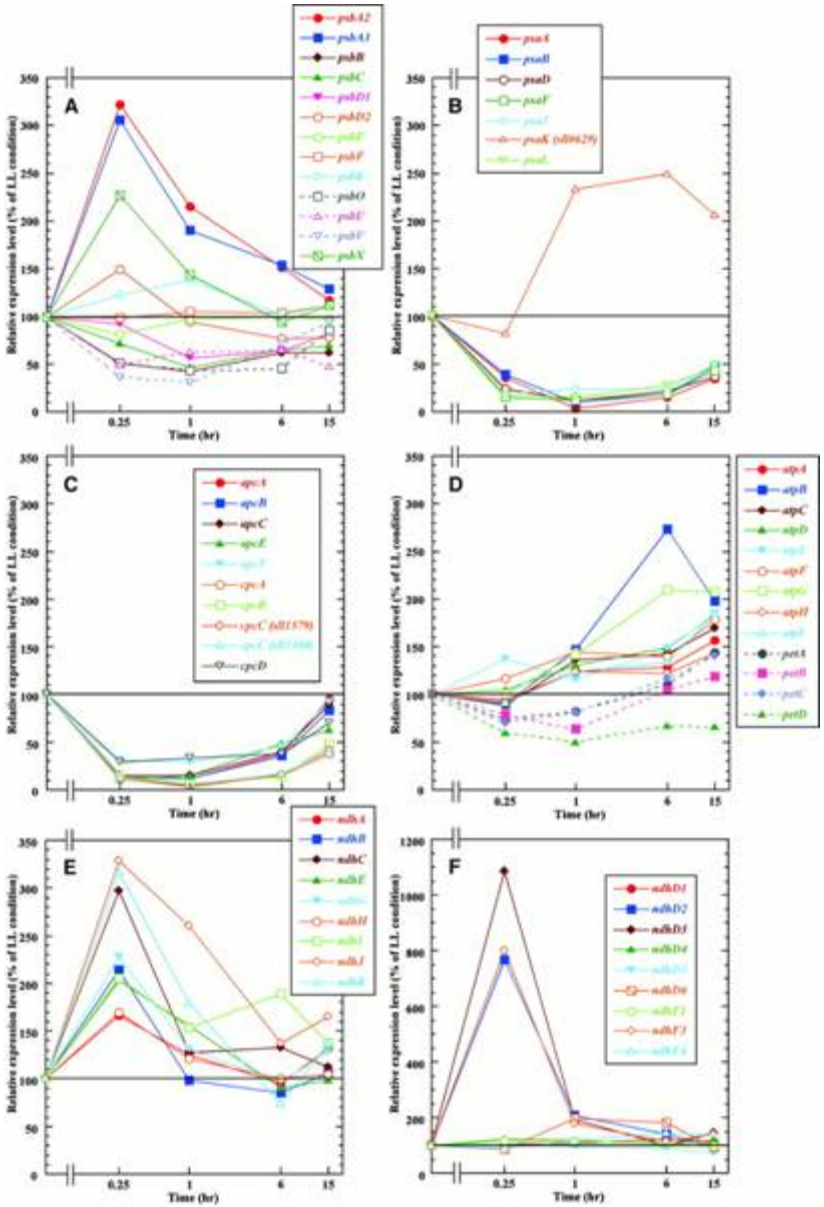


Prochlorococcus MIT9313

Hihara, Kamei, Kanehisa, Kaplan, and Ikeuchi (2001) DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. Plant Cell, 13(4)



Synechocystis PCC 6803



How do cells control response to light?

I.e., What genes are related to the adaptation to high light?

Outline Protocol

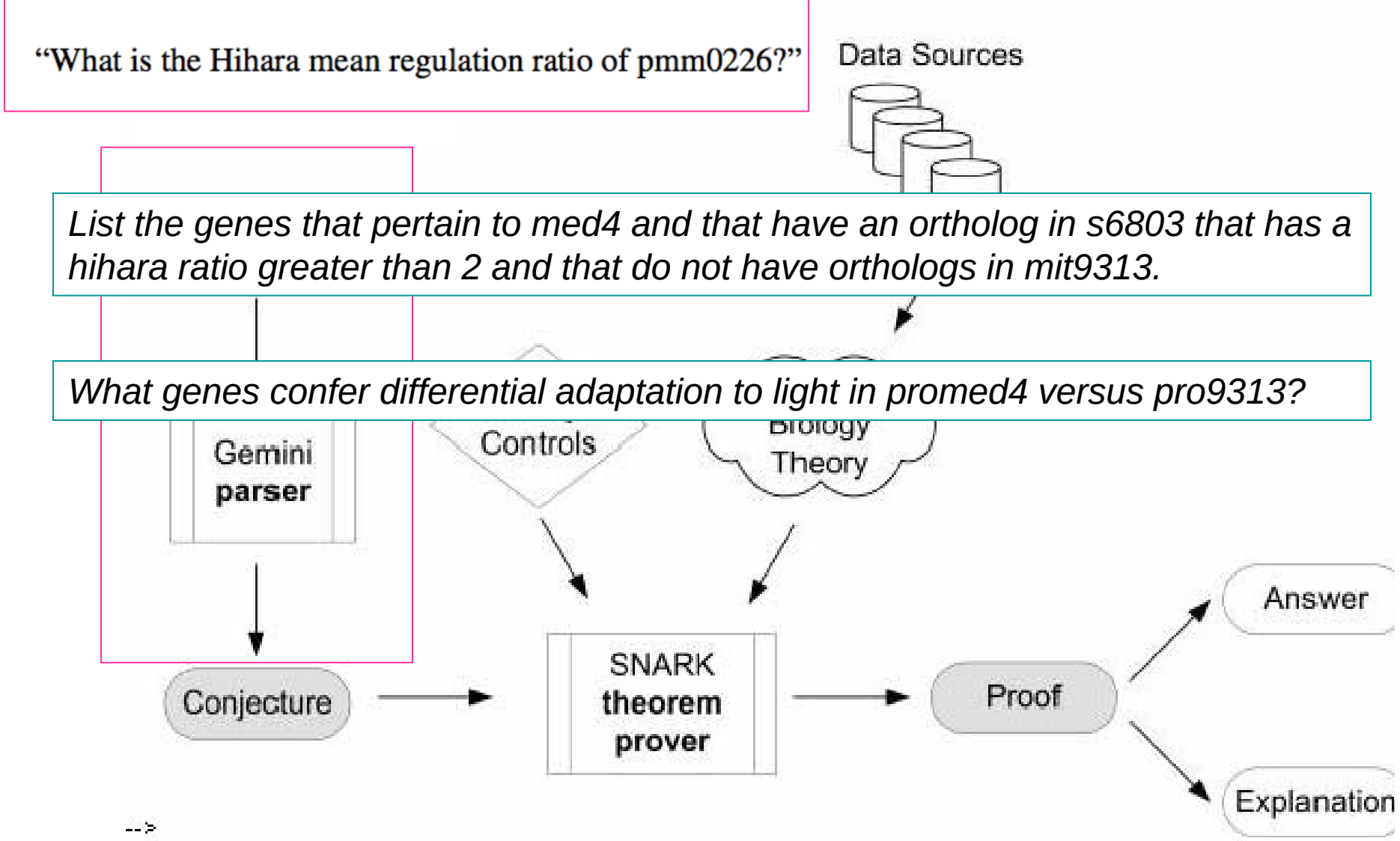
Look for:

- Gene present in *Prochlorococcus* MED4
  MED4 is naturally adapted to grow in high light.
- Ortholog absent in *Prochlorococcus* MIT9313
  MIT9313 is naturally adapted to grow in low light
- Ortholog present in *Synechocystis* PCC 6803
  In order to make contact with annotation and microarray data
- *Synechocystis* PCC 6803 ortholog responds to high light
  Gene turns on by factor > 2 in response to high light

Natural Language Deductive Biocomputing

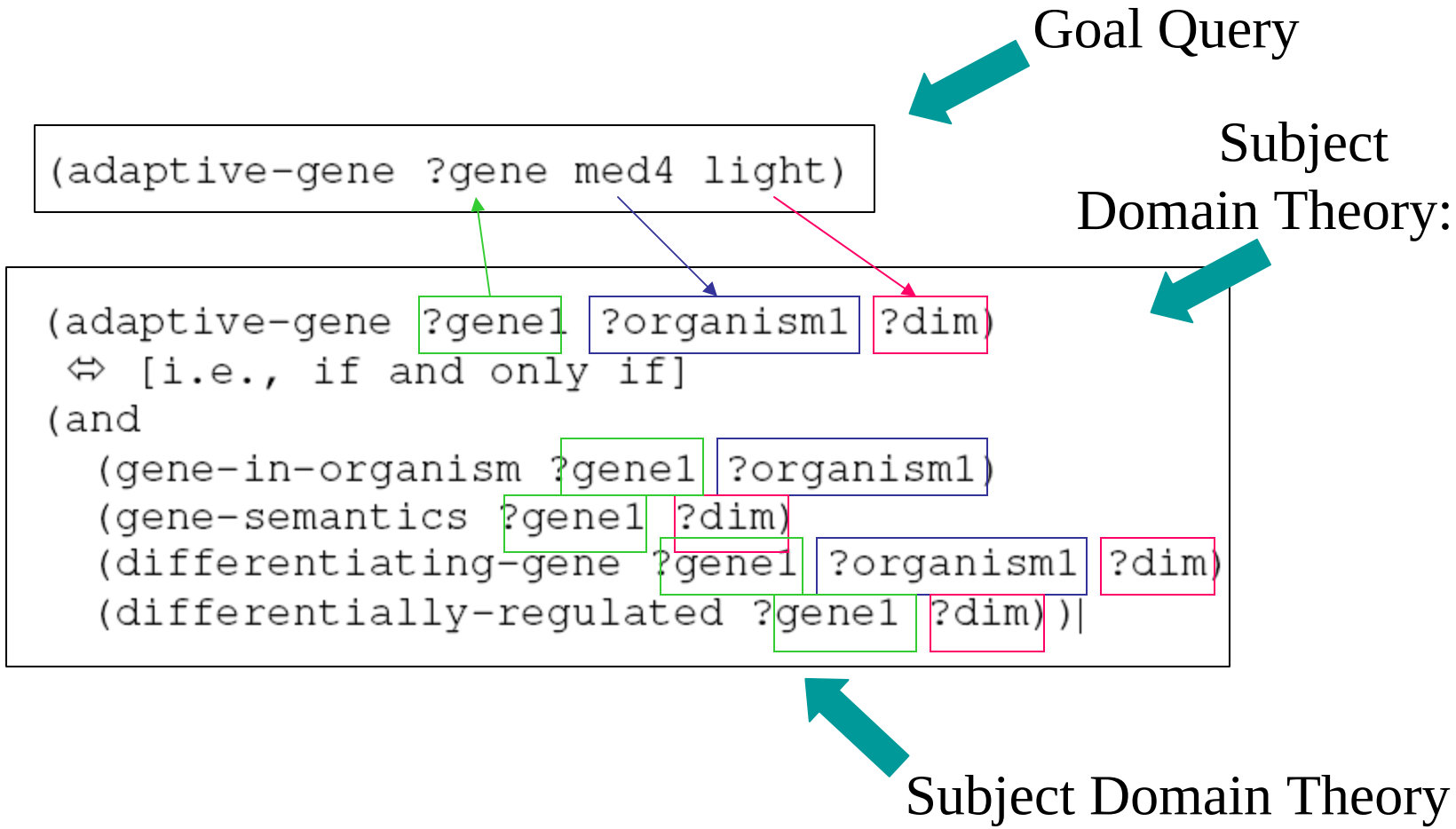
"Does pmm0226 not have an ortholog in mit9313?"

"What is the Hihara mean regulation ratio of pmm0226?"



Language for Expressing Conjectures, and Platform for Analysis

A. First Order Logic (FOL) representation
B. Subject Domain Theory
C. Biological Process (and entities) Ontology
D. Visual query language.



```
(differentially-regulated ?gene ?dim)

⇔
(exists
  ((?experiment :sort experiment)
   (?organism3 :sort species)
   (?gene3 :sort gene))
 (and
   (experiment ?experiment :dimension ?dim
                           :organism ?organism3)
   (gene-has-ortholog-in-organism ?gene
                                  ?gene3 ?organism3)
   (> (regulation-ratio ?gene3 ?experiment) 2.0)
   ))
```

```
(experiment hihara :dimension light :organism s6803)
```

Goal Query: `(adaptive-gene ?gene med4 light)`

Result:

?gene: #$PMED4.PMM0817
?organism2: #$prochlorococcus_marinus_mit9313
?experiment: HIHARA
?organism3: #$synechocystis_pcc6803
?gene3: #$S6803.ssr2595

I.e., A low-light organism that has no ortholog to ?gene is prochlorococcus marinus pcc. 9313. Experiments were performed by Hihara on the organism synechocystis pcc 6803, and a high regulation ratio was discovered in those experiments on gene S6803.ssr2595, which is an ortholog of PMM0817. The annotation for PMM0817 reads: “possible high-light inducible protein”.

(Matches the results from: Bhaya, Dufresne, Vaulot, and Grossman: Analysis of the hli gene family in marine and freshwater cyanobacteria. FEMS Letters, 2002, 205(2). PMM0817 is called hli17 in this paper.)

Goal Query: (adaptive-gen

Result:

?gene: #$PMED4.PMM081
?organism2: #$prochlo
?experiment: HIHARA
?organism3: #$syntecho
?gene3: #$S6803.ssr25

+ "Explanation"

```
   (resolve 5411 definition-of-differentiating-gene)
   Answer (or (snark::answer-- (list |hashdollar-PMED4.PMM0817|))
              (defn-of-differentially-regulated
               |hashdollar-PMED4.PMM0817| light
               |hashdollar-S6803.ssr2595|
               |hashdollar-synechocystis_pcc6803|)))
(Row 5416
   (or (not (organism ?species :environment light :quality low))
    (not
     (gene-in-organism |hashdollar-PMED4.PMM0817|
      |hashdollar-prochlorococcus_marinus_med4|))
    (gene-has-ortholog-in-organism |hashdollar-PMED4.PMM0817|
     (snark-user::gene-skolemmqlk2 |hashdollar-PMED4.PMM0817| ?species)
     ?species))
   (rewrite (resolve 5413 defn-of-differentially-ecotyped) 8)
   Answer (or (snark::answer-- (list |hashdollar-PMED4.PMM0817|))
              (defn-of-differentially-regulated
               |hashdollar-PMED4.PMM0817| light
               |hashdollar-S6803.ssr2595|
               |hashdollar-synechocystis_pcc6803|)))
(Row 5422
   (or
    (not
     (gene-in-organism |hashdollar-PMED4.PMM0817|
      |hashdollar-prochlorococcus_marinus_med4|))
    (gene-has-ortholog-in-organism |hashdollar-PMED4.PMM0817|
     (snark-user::gene-skolemmqlk2 |hashdollar-PMED4.PMM0817|
      |hashdollar-prochlorococcus_marinus_mit9313|)
     |hashdollar-prochlorococcus_marinus_mit9313|))
   (resolve 5416 7)
   Answer (or (snark::answer-- (list |hashdollar-PMED4.PMM0817|))
              (defn-of-differentially-regulated
               |hashdollar-PMED4.PMM0817| light
               |hashdollar-S6803.ssr2595|
               |hashdollar-synechocystis_pcc6803|)))
(Row 5425
   (or
    (not
     (gene-in-organism-code |hashdollar-PMED4.PMM0817|
      |hashdollar-prochlorococcus_marinus_med4|))
    (gene-has-ortholog-in-organism |hashdollar-PMED4.PMM0817|
     (snark-user::gene-skolemmqlk2 |hashdollar-PMED4.PMM0817|
      |hashdollar-prochlorococcus_marinus_mit9313|)
     |hashdollar-prochlorococcus_marinus_mit9313|))
   (resolve 5422 definition-of-gene-in-organism-code)
   Answer (or (snark::answer-- (list |hashdollar-PMED4.PMM0817|))
              (defn-of-differentially-regulated
               |hashdollar-PMED4.PMM0817| light
               |hashdollar-S6803.ssr2595|
               |hashdollar-synechocystis_pcc6803|)))
(Row 5427
   (gene-has-ortholog-in-organism |hashdollar-PMED4.PMM0817|
    (snark-user::gene-skolemmqlk2 |hashdollar-PMED4.PMM0817|
```

How Science Thinks:
The Science and Engineering of Science and Engineering

For every right answer in science,
There is an infintitude(-1) of wrong ones!

HOW COULD THIS EVER WORK!?

1. Close is often good enough, or at least guides you to the right answer.

2. Theory (model) guidance reduces the search space by huge orders!

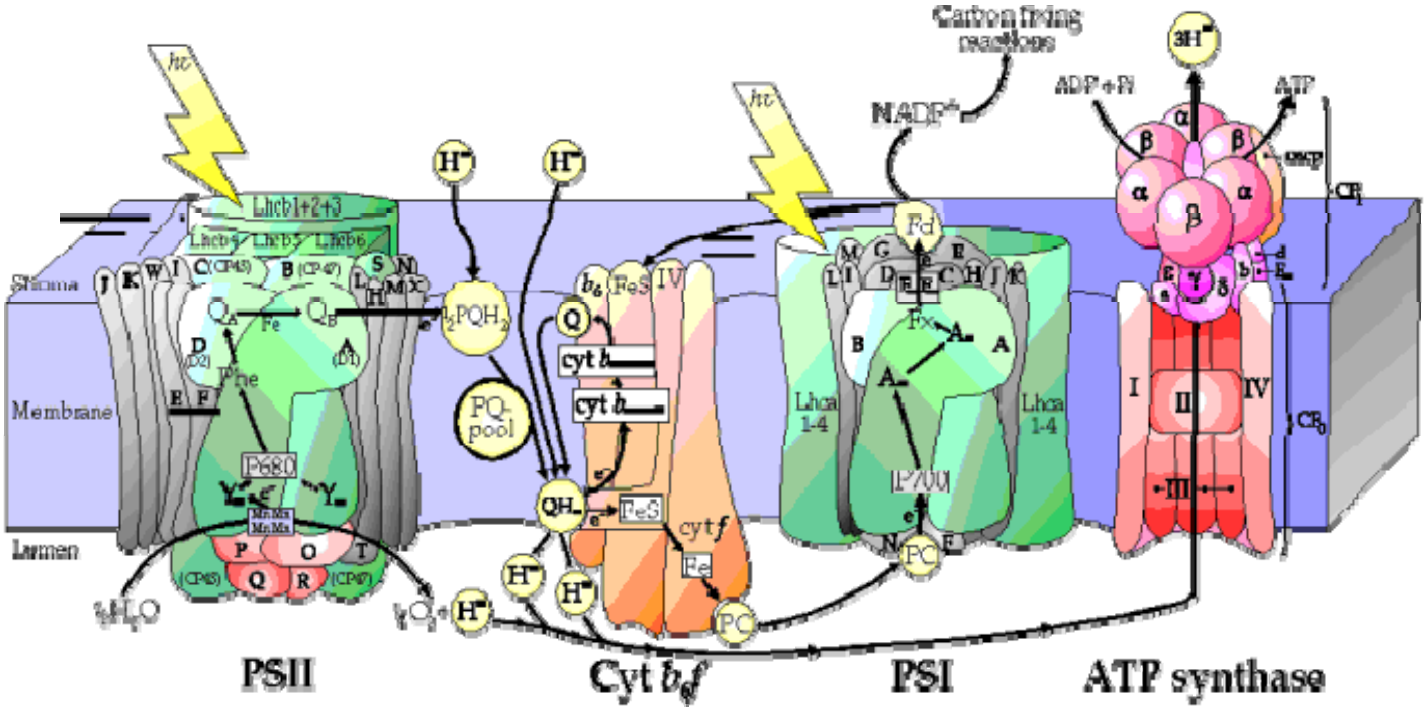
3. We've been really *really* lucky ... so far, anyway!

4. *You are not alone!* (>15Million abstract in pubmed alone!)

Simulation: What does this model predict?

Explanation: How does it make these predictions?

Model Identification: What models fit this data?

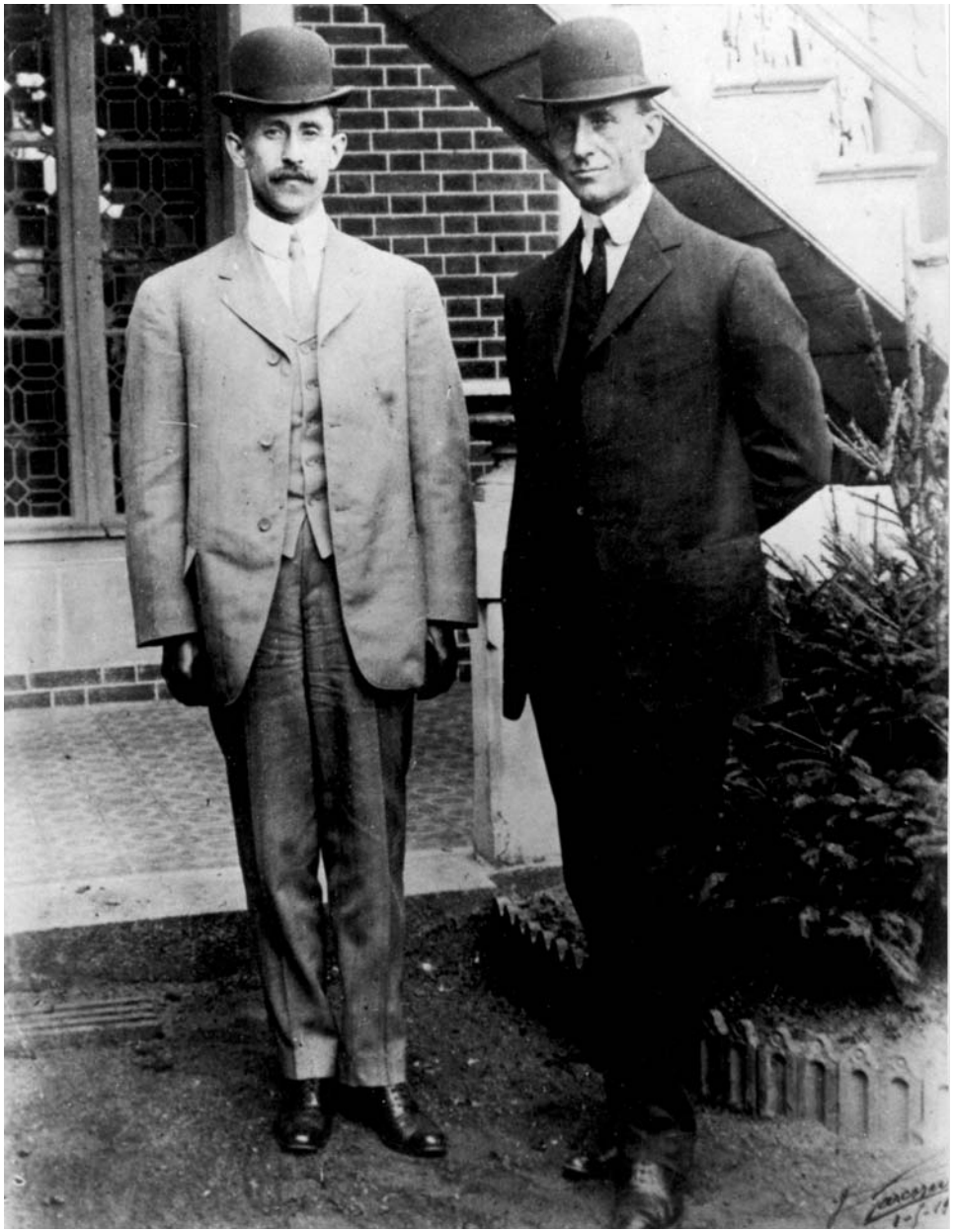


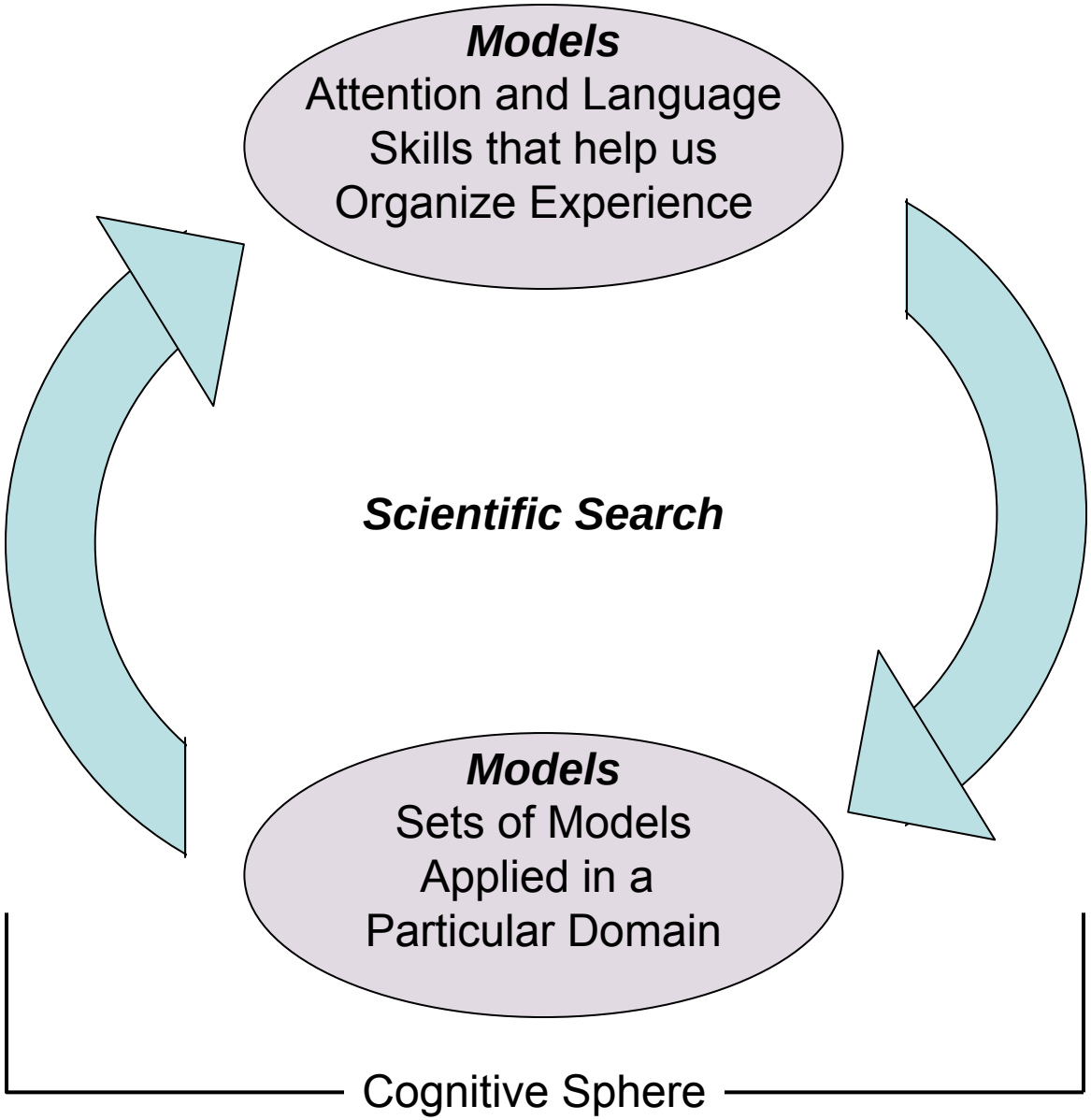
Simulation: What does this model predict?

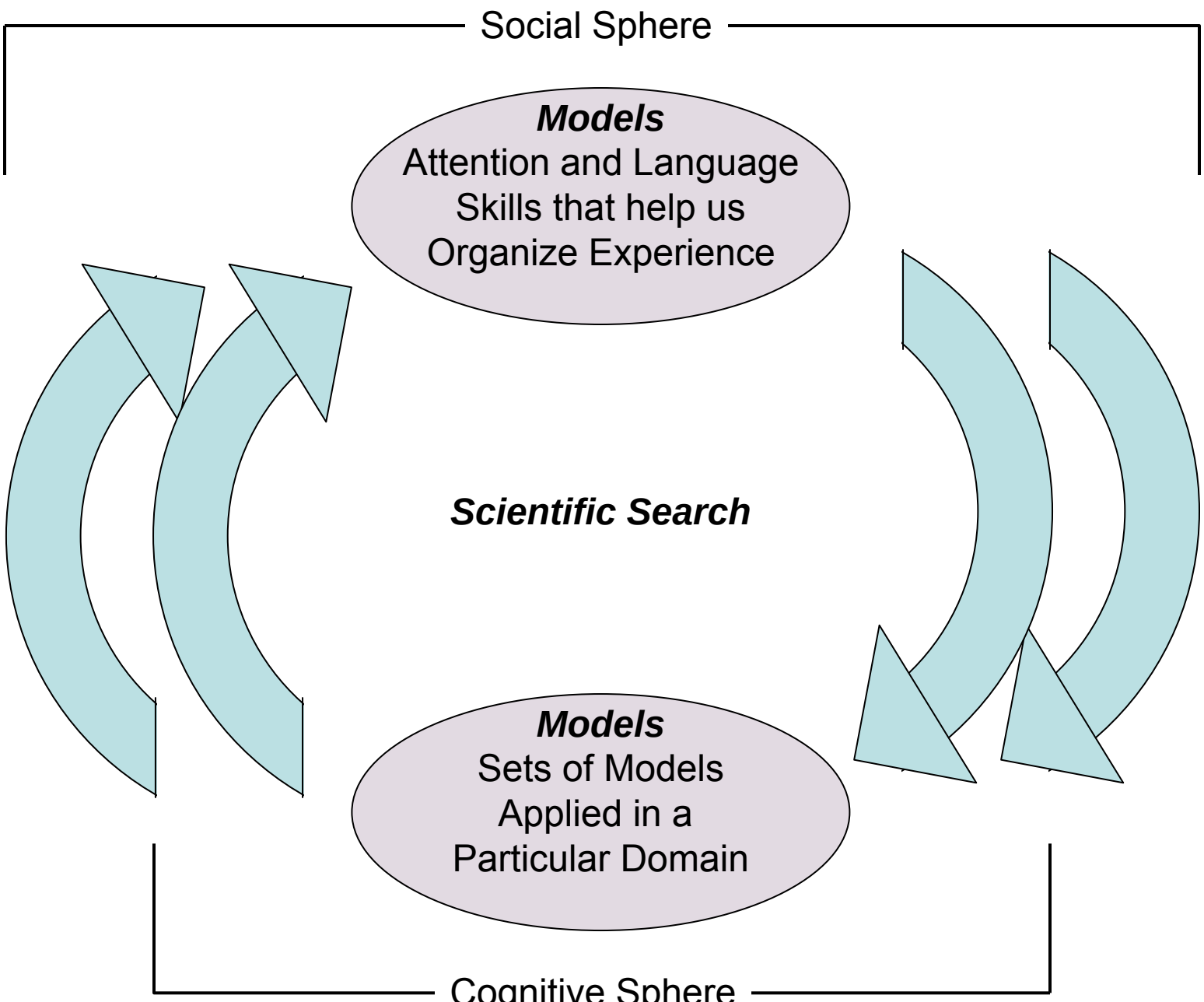
Explanation: How does it make these predictions?

Model Identification: What models fit this data?

Collaboration: Interact with scientists...
...and help scientists interact with one another!




Models
Attention and Language
Skills that help us
Organize Experience
Scientific Search
Models
Sets of Models
Applied in a
Particular Domain
Cognitive Sphere


Social Sphere
Models
Attention and Language
Skills that help us
Organize Experience
Scientific Search
Models
Sets of Models
Applied in a
Particular Domain
Cognitive Sphere

Historical Theory↔Experiment SeeSawing of the Gyromagnetic Ratio “g-factor”

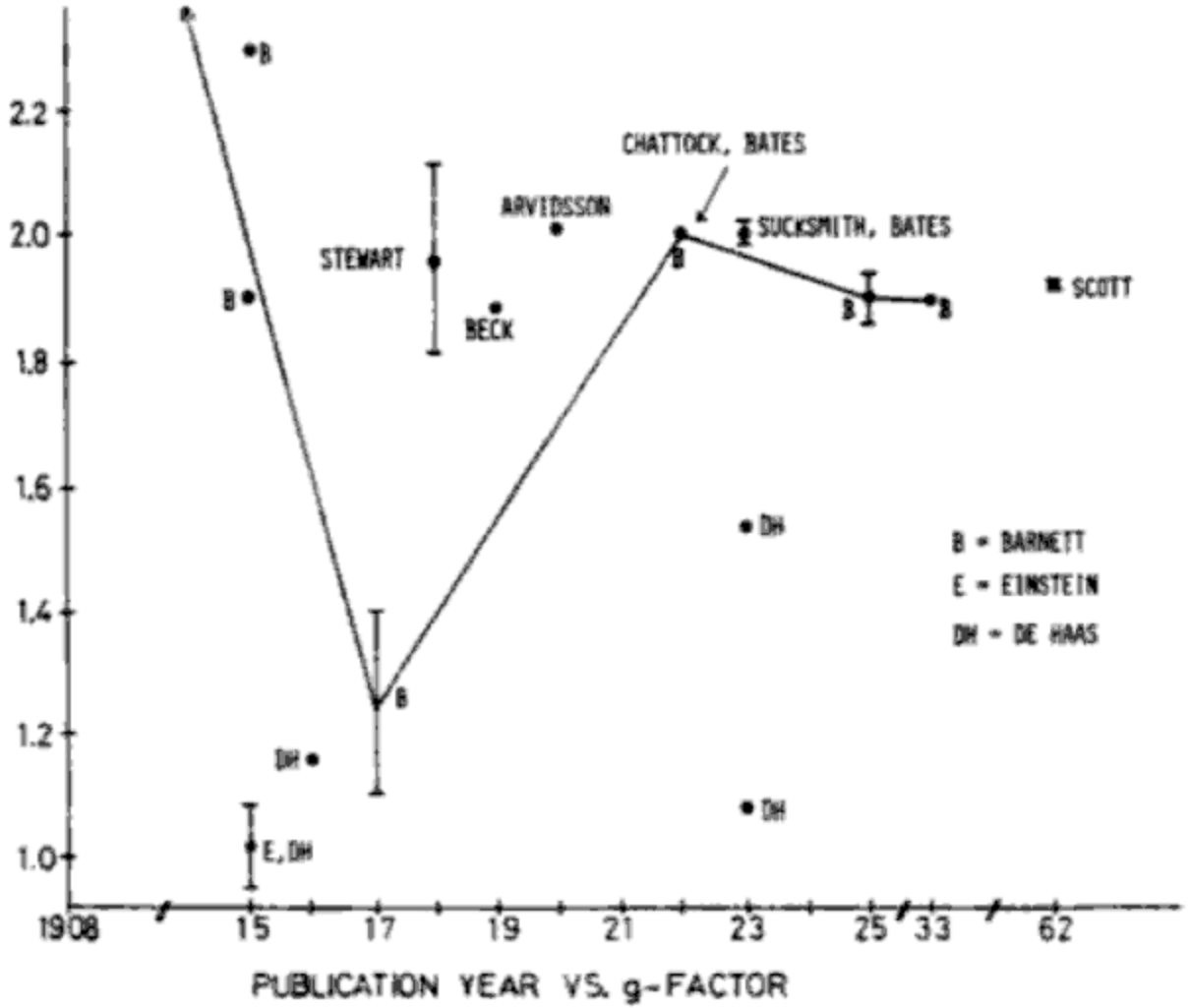
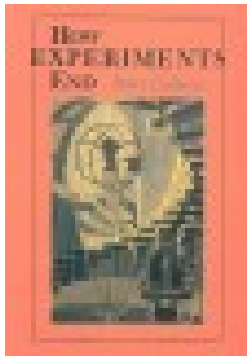


FIGURE 2.15. Publication year versus g-factor. The solid line traces Barnett's results as a function of time. Einstein's theoretical prediction was that $g = 1$; later experiment and theory both gave $g \sim 2$.



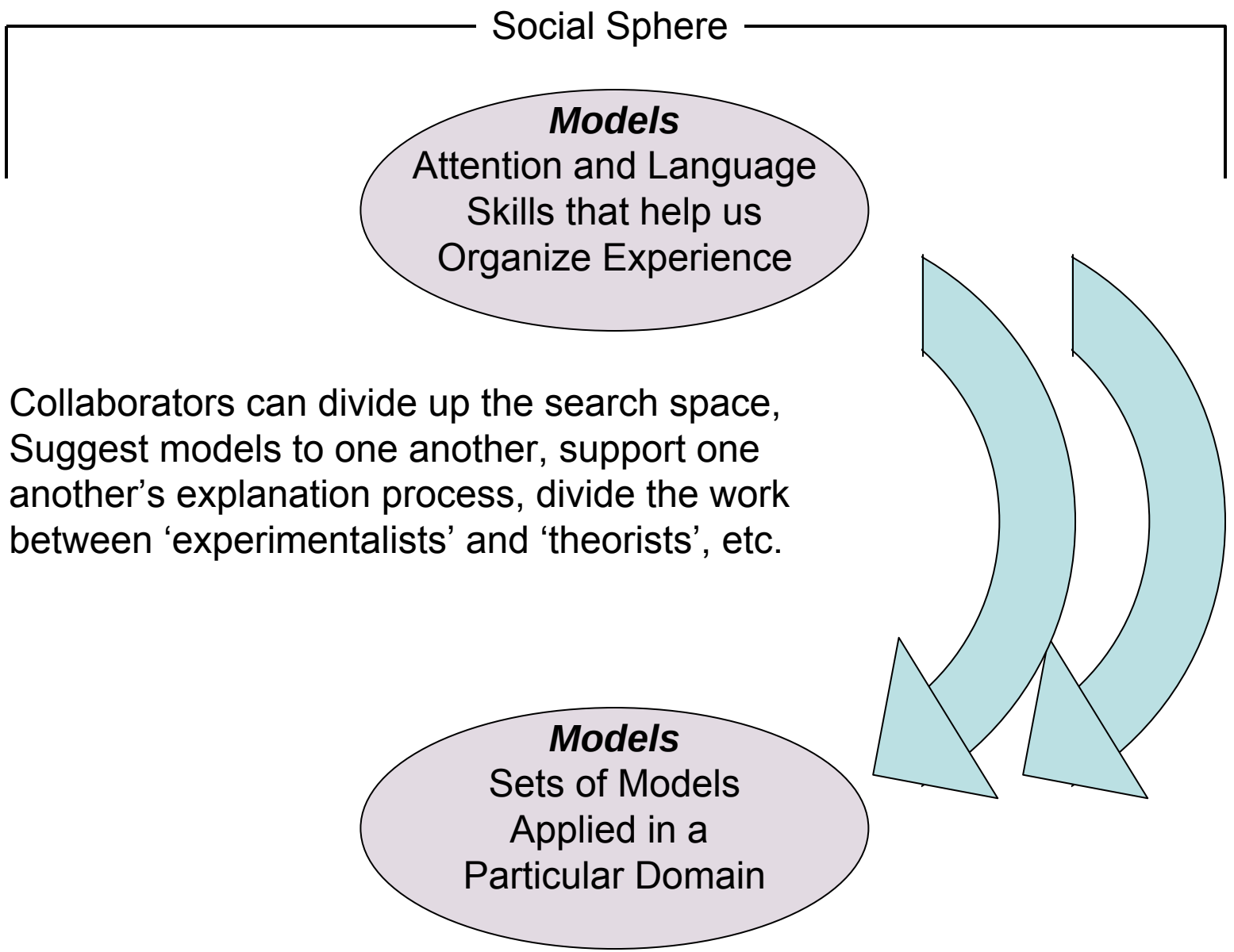
Galison (1987) How Experiments End. Chicago U. Press

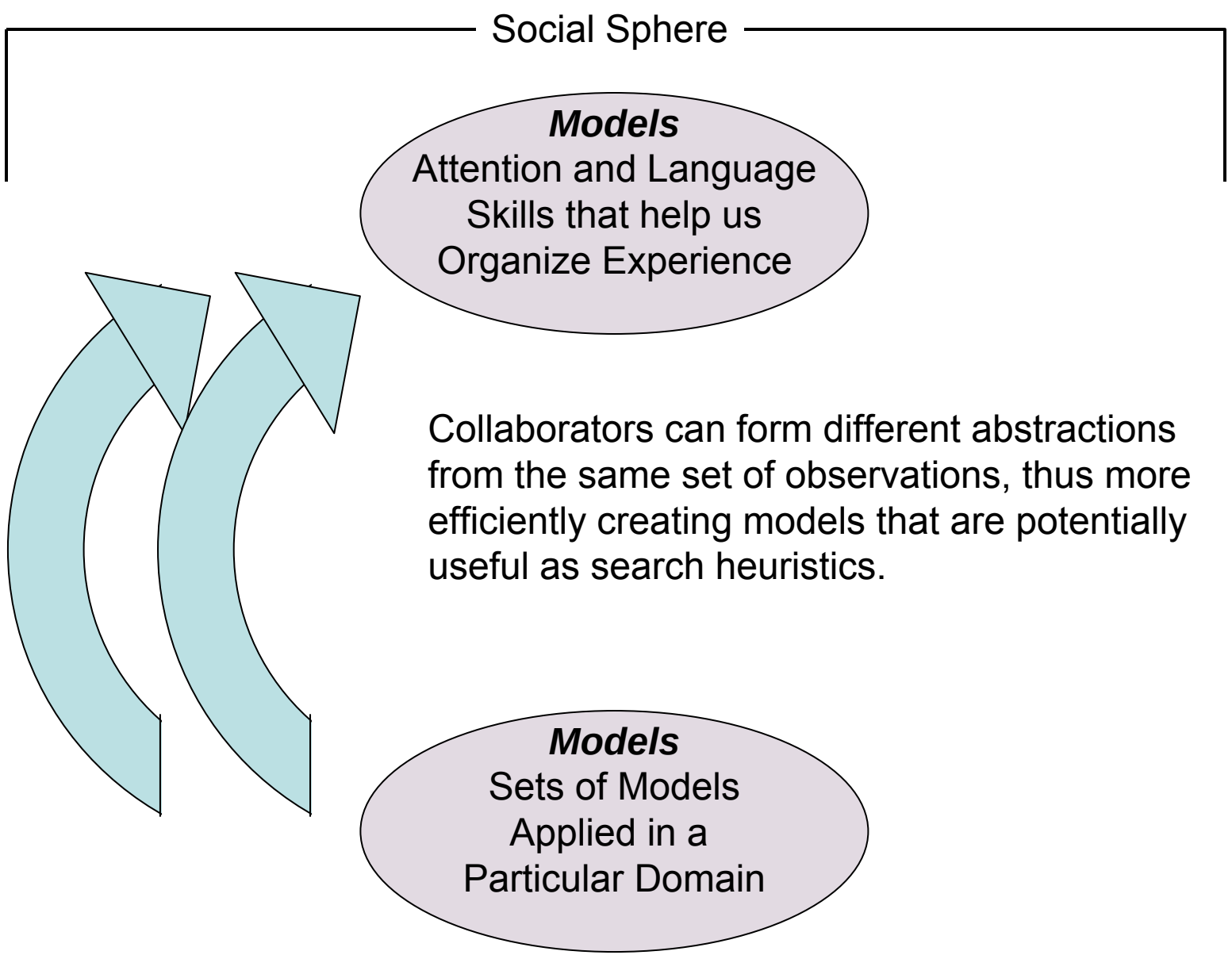
ory that orbiting electrons are responsible for permanent magnetism. They determined that $g = 1.02 \pm 0.10$; and, in a second quantitative series of experiments the following year, de Haas found $g = 1.2$. Then Barnett, obviously influenced by Einstein's theory and experiment, repeated his own work and concluded that he too had vindicated the orbiting electron theory: g was somewhere between 1.4 and 1.1.

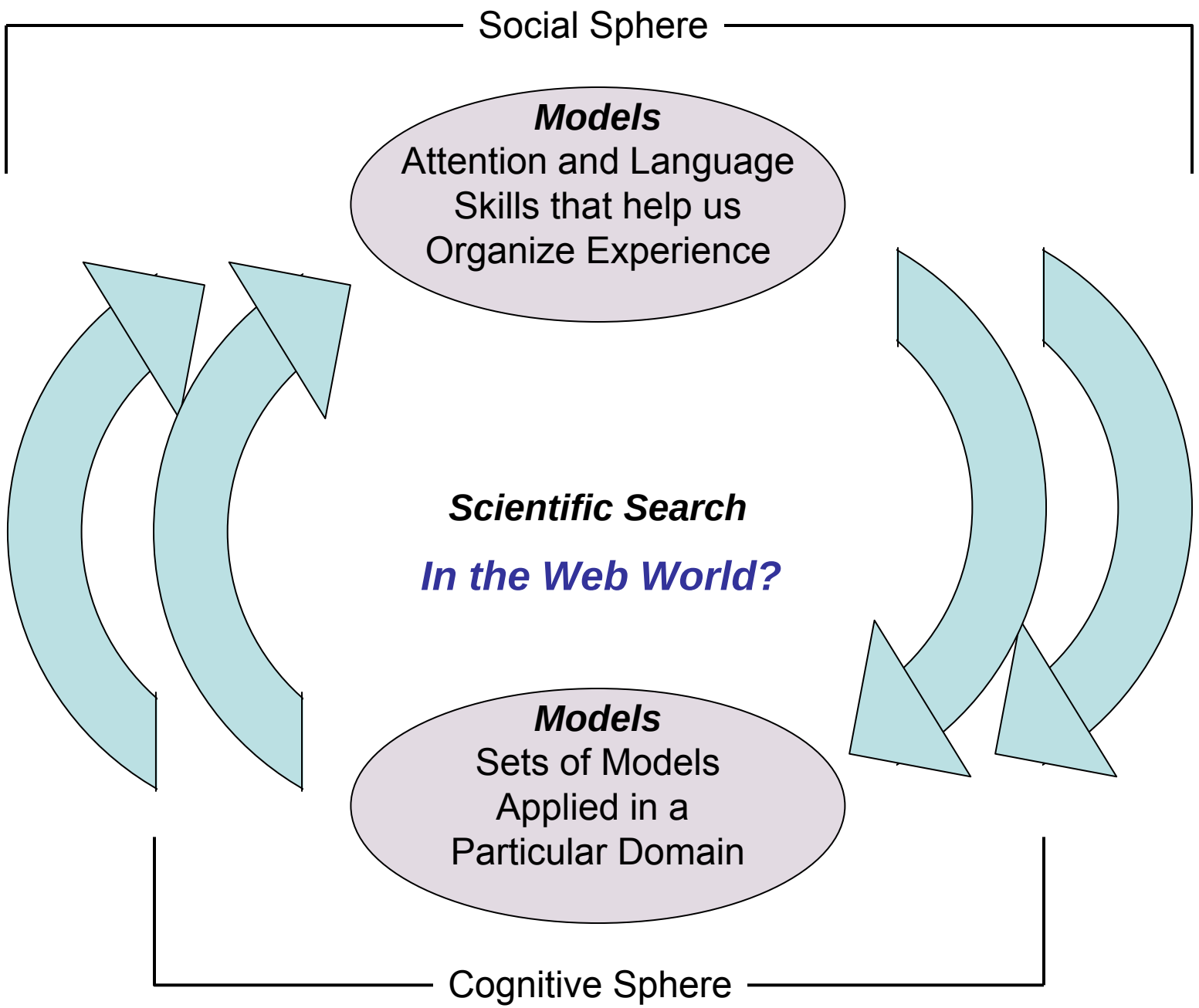
Three experimentalists, working independently, soon determined that g was *not* equal to 1. Stewart, Beck, and Arvidsson each published a tentative result nearer to twice the Einstein value. Within months Barnett published again, asserting that he too believed that g was approximately 2. In the two years that followed, he improved his result, abandoned Einstein's theory, and adopted one of Abraham's electron theories to explain his result of $g = 1.89$.

Meanwhile, de Haas (1921) repeated his work, now aware that at least four other researchers were finding a g-value nearly twice his original one. At the Solvay meeting of 1921, he reported a g-value of 1.54 and said that he still considered the value of g to be an open question. Afterward he repeated his experiments for the last time: his cumulative result of $g = 1.08$ was only a few percent different from his original result with Einstein six years earlier. The next year, in Berlin, Einstein too maintained that the value of g was still open to question. During this time Barnett and his wife refined their method further and in 1925 published a massive paper with an average result of $g = 1.929 \pm 0.006$. At first

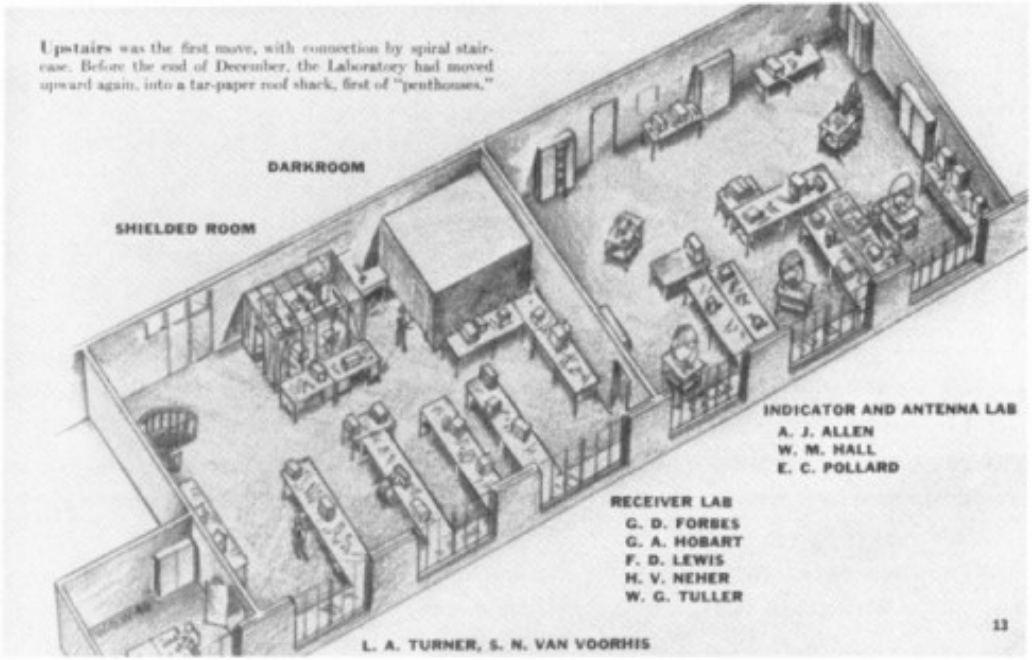
Galison (1987) How Experiments End. Chicago U. Press


Social Sphere
Models
Attention and Language
Skills that help us
Organize Experience
Collaborators can divide up the search space,
Suggest models to one another, support one
another's explanation process, divide the work
between 'experimentalists' and 'theorists', etc.
Models
Sets of Models
Applied in a
Particular Domain

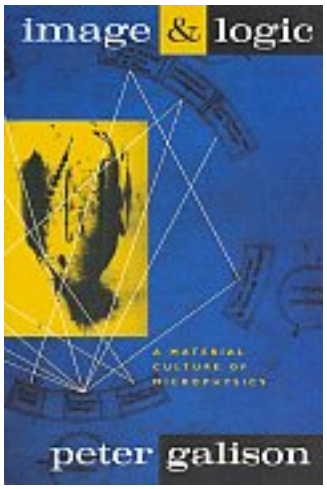

Social Sphere
Models
Attention and Language
Skills that help us
Organize Experience
Collaborators can form different abstractions from the same set of observations, thus more efficiently creating models that are potentially useful as search heuristics.
Models
Sets of Models
Applied in a
Particular Domain


Social Sphere
Models
Attention and Language
Skills that help us
Organize Experience
Scientific Search
In the Web World?
Models
Sets of Models
Applied in a
Particular Domain
Cognitive Sphere

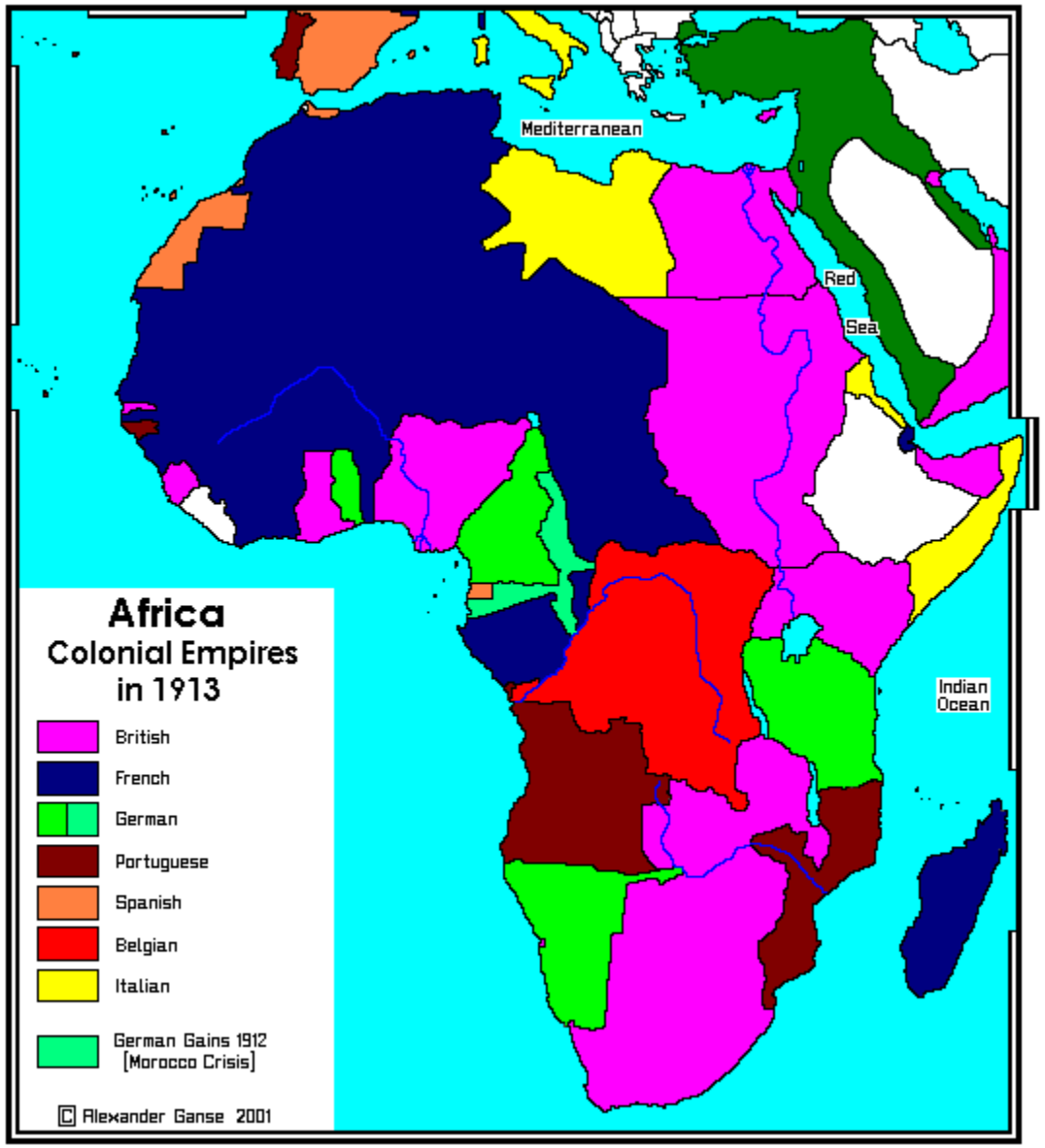
Scientific Collaborations as “Trading Zones”



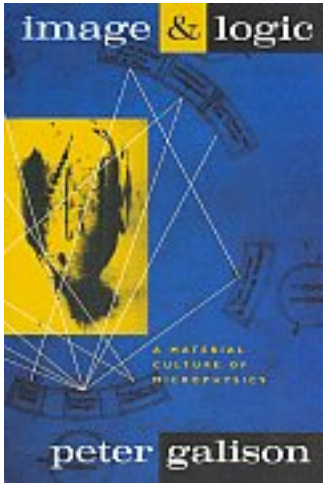
Galison, Image and Logic, p.819



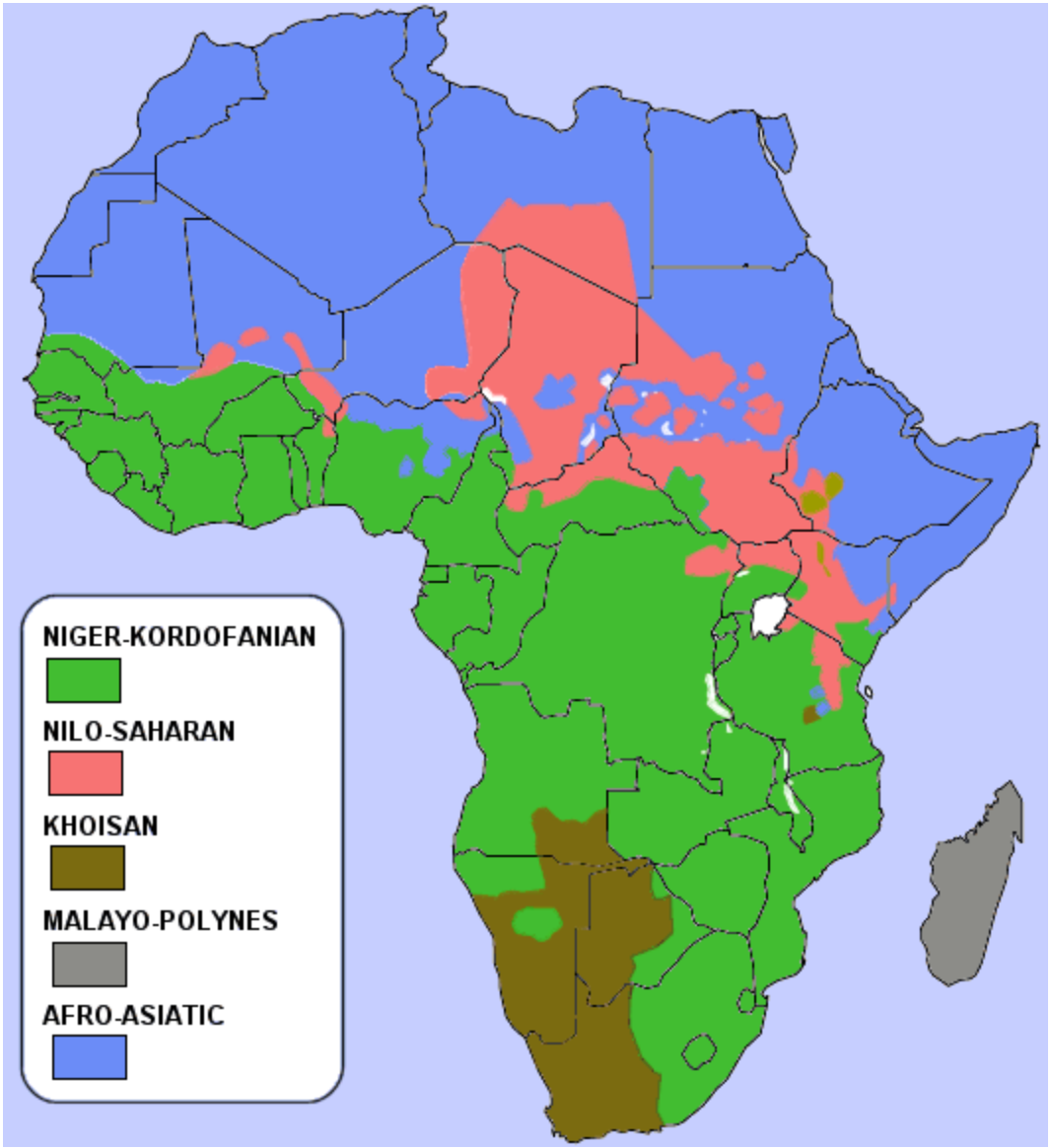
Scientific Collaborations as “Trading Zones”



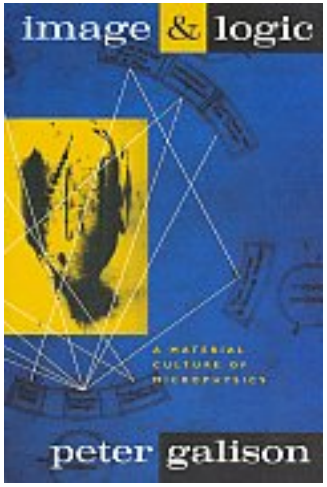
http://www.zum.de/whkmla/histatlas/africa/colafr1913.gif



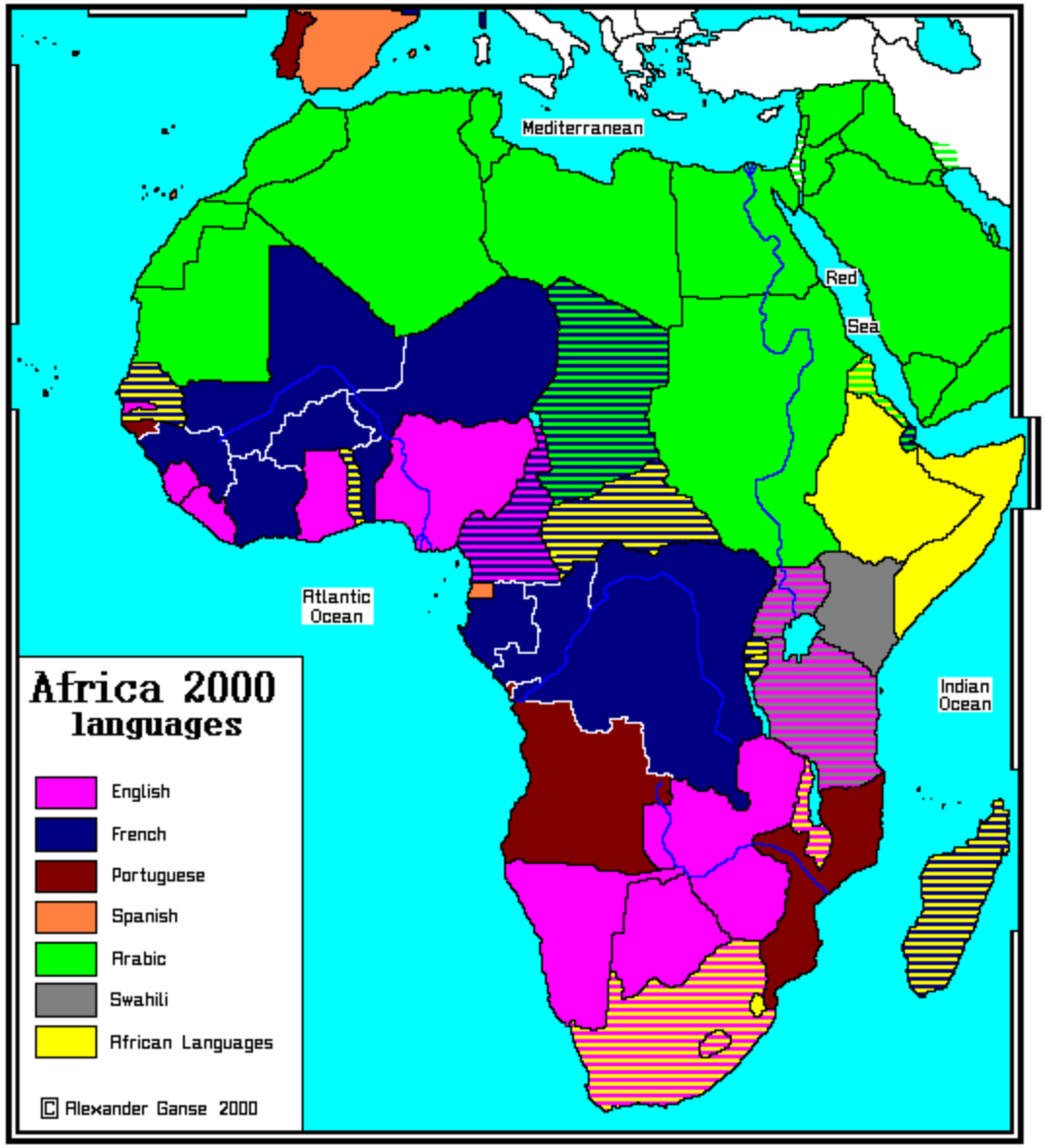
Scientific Collaborations as “Trading Zones”



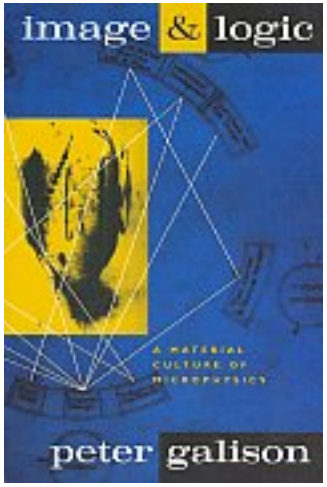
http://www.artsci.wustl.edu/~anthro/courses/306/africa_linguistic_map.gif



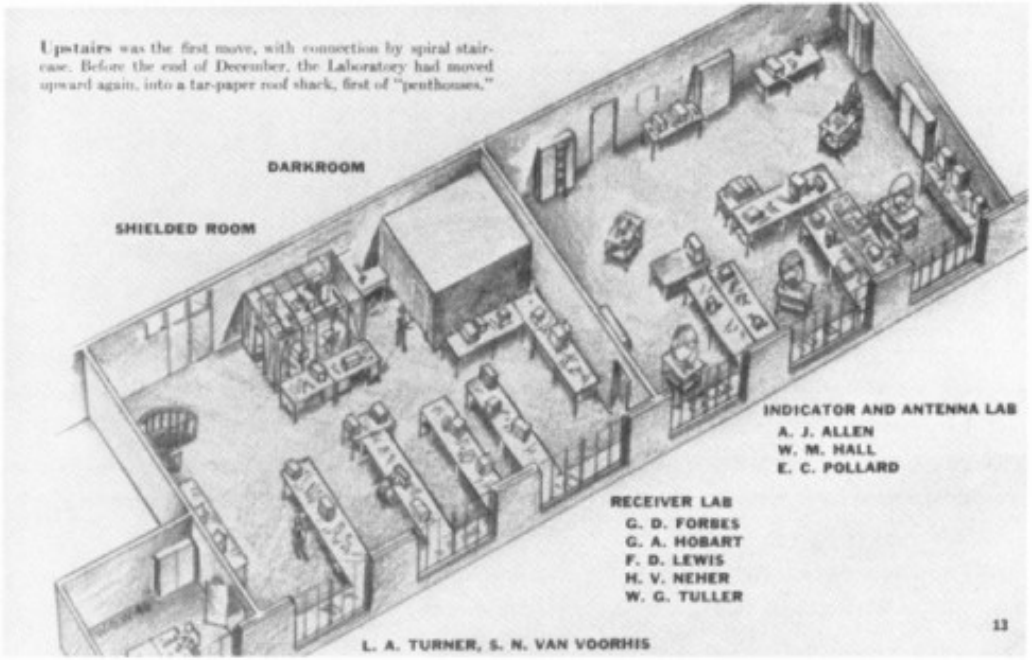
Scientific Collaborations as “Trading Zones”



www.zum.de/whkmla/histatlas/africa/afr95lang.gif

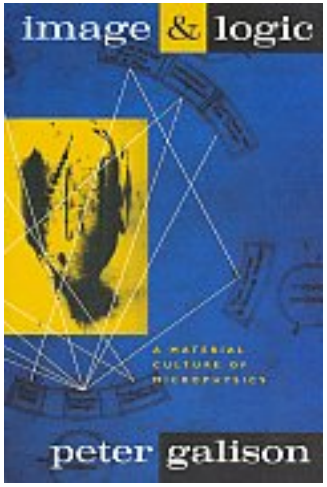


Scientific Collaborations as “Trading Zones”

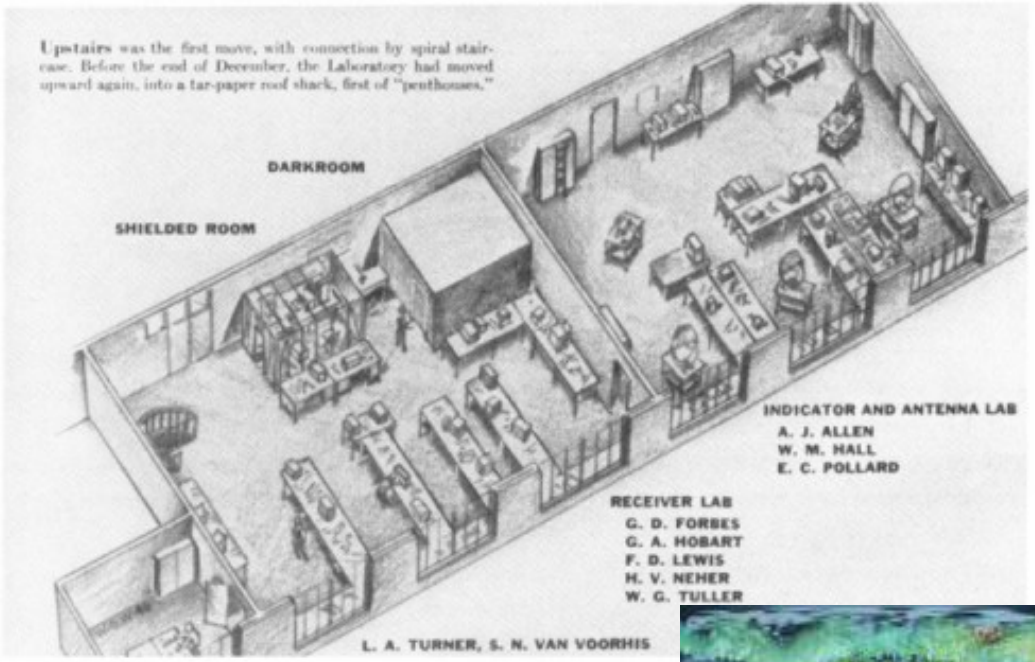


“...engineers structured their work around components, rather than ... Around ‘pure’ and ‘applied’ science. Working out a common language became the order of the day.”

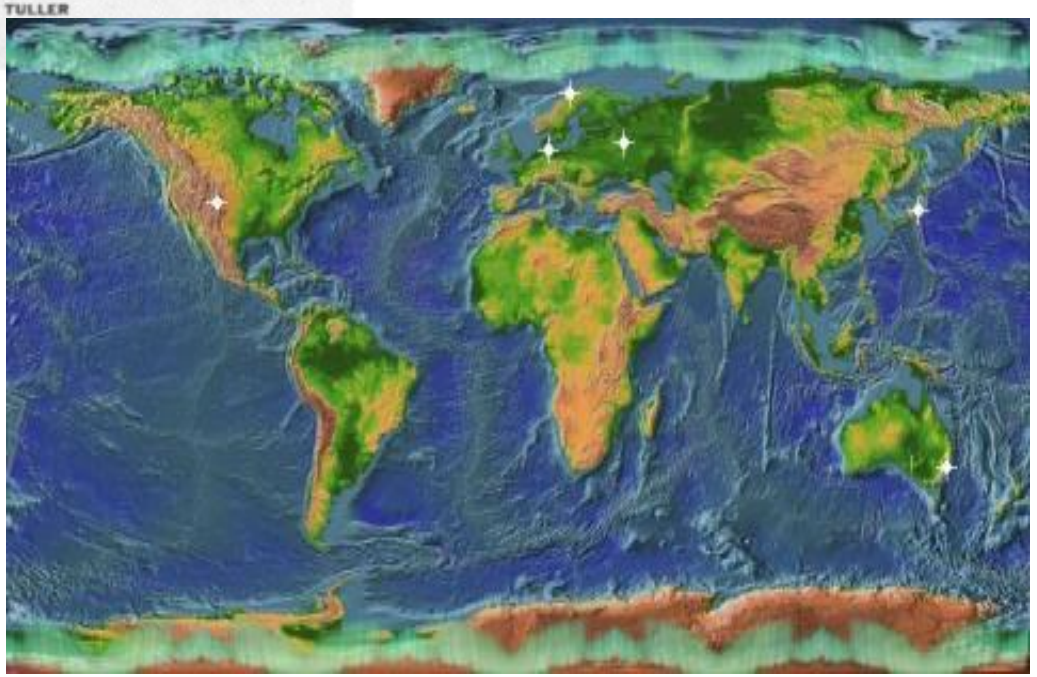
Galison, Image and Logic, p.819

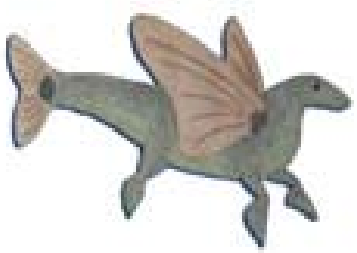


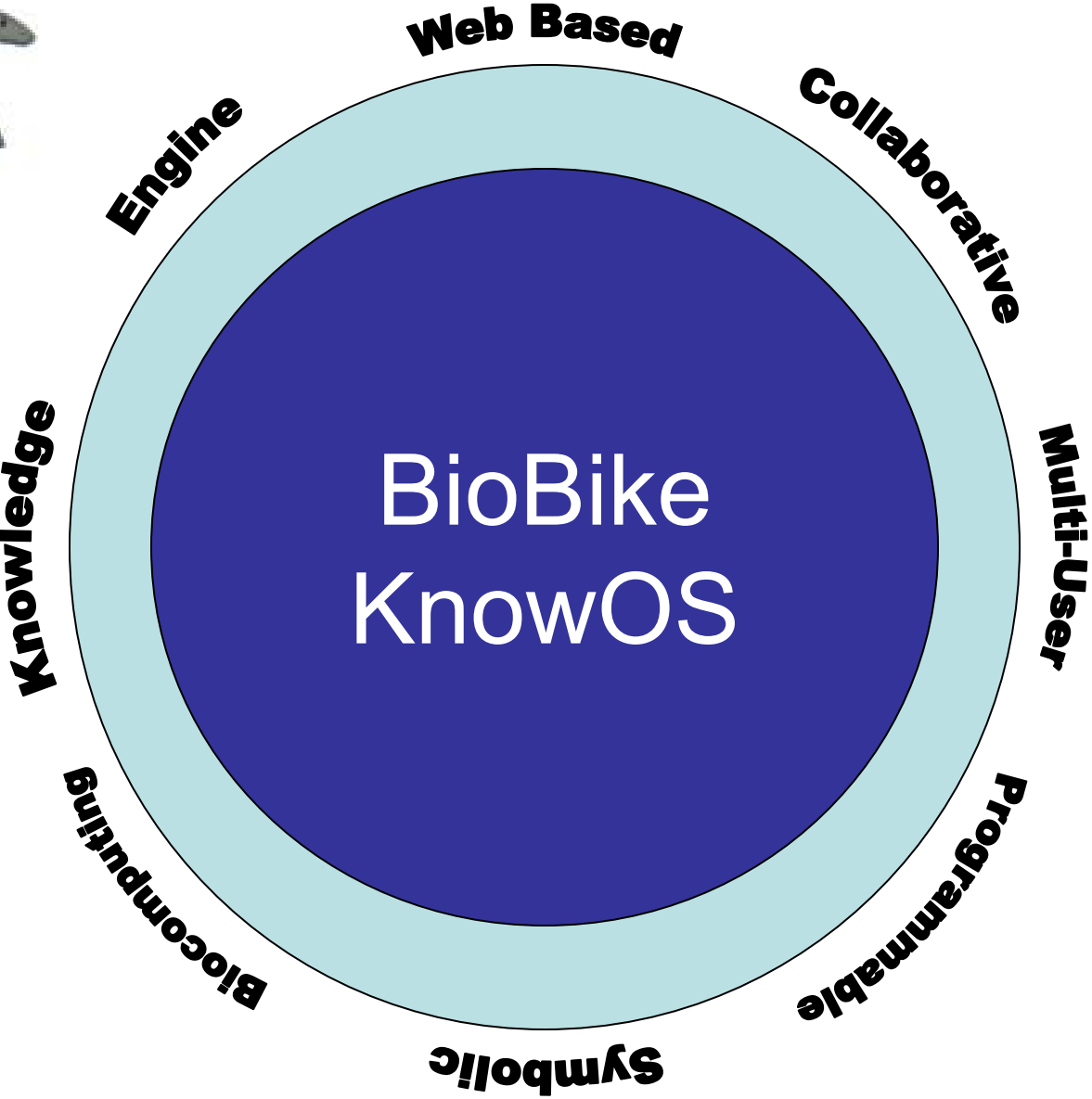
Scientific Collaborations as “Trading Zones”



In the Web World?

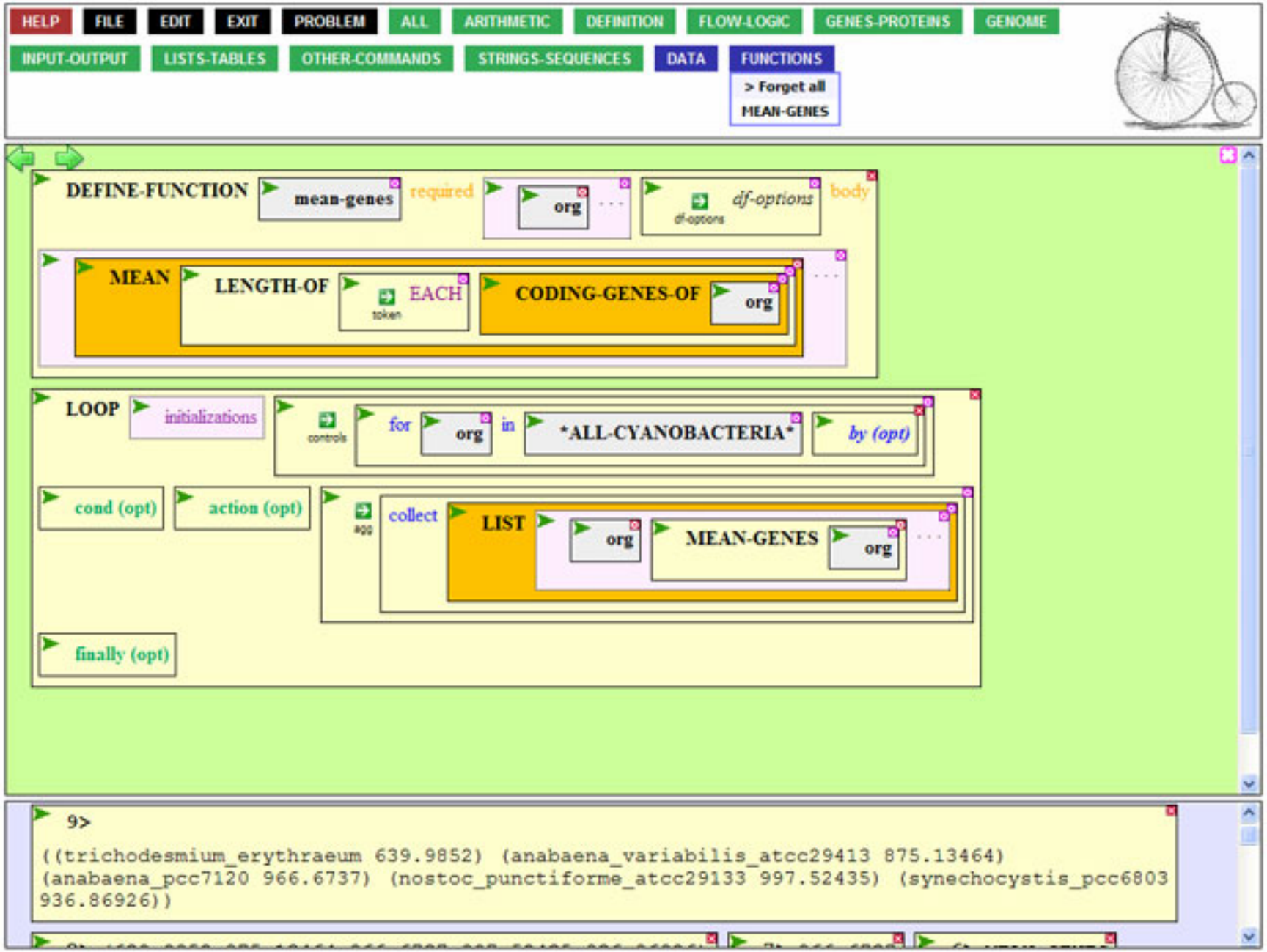




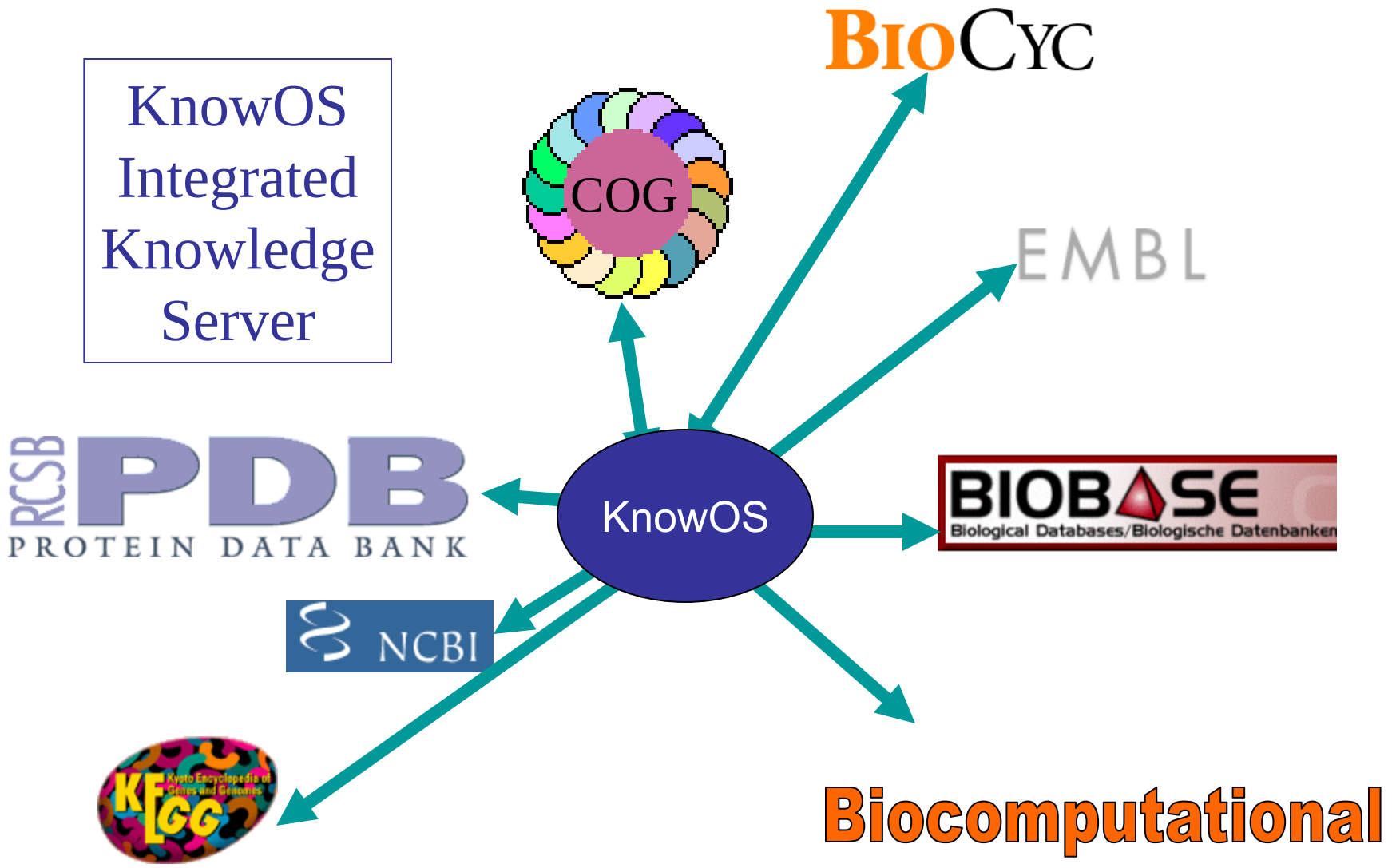


www.biobike.org

www.knowos.org


HELP
FILE
EDIT
EXIT
PROBLEM
ALL
ARITHMETIC
DEFINITION
FLOW-LOGIC
GENES-PROTEINS
GENOME
INPUT-OUTPUT
LISTS-TABLES
OTHER-COMMANDS
STRINGS-SEQUENCES
DATA
FUNCTIONS
> Forget all
MEAN-GENES
DEFINE-FUNCTION
mean-genes
required
org
df-options
body
MEAN
LENGTH-OF
EACH
token
CODING-GENES-OF
org
LOOP
initializations
controls
for
org
in
ALL-CYANOBACTERIA
by (opt)
cond (opt)
action (opt)
collect
LIST
org
MEAN-GENES
org
finally (opt)
9>
((trichodesmium_erythraeum 639.9852) (anabaena_variabilis_atcc29413 875.13464)
(anabaena_pcc7120 966.6737) (nostoc_punctiforme_atcc29133 997.52435) (synechocystis_pcc6803
936.86926))

Knowledge Operating System Integrates Knowledge Resources



BioBike/KnowOS Integrates Scientists and Computation in a *Trading Zone*

KnowOS Integrated Knowledge Server

BIOCYC

RCSB PDB PROTEIN DATA BANK

KnowOS

BIOBASE Biological Databases/Biologische Datenbanken

Biocomputational Tools and Sims

BioBike/KnowOS is a "Web 24.0" Platform:

Web 1.0: The "Page" Web
Web 2.0: The "Social" Web
Web 3.0: The "Semantic" Web
Web 4.0: The "Programmable" Web

1.0 x 2.0 x 3.0 x 4.0 = 24.0000000000001:
The "*Social Semantic Programmable*" Web!

From: DR. X <[Michigan]>
Date: Oct 21, 2004 7:09 AM
Subject: Help with BioLingua

I'm a new user of BioLingua, with very little experience in computer programming. I'm searching for housekeeping genes in Anabaena 7120 that are longer than 3000 bp. I could load Anabaena sequences by:

>> (setf an (load-organism "A7120"))

I found genes that are involved in metabolism by:

>> (setf metabolism (find-frames "metabolism"))

I got a list of related genes by:

>> (df #$go.metabolism)

now I want to find the length of each gene in the list "metabolism" and check if it is longer than 3000. This is where I don't know what function to use.

I tried to start with the loop:

```
(LOOP FOR LongSequences in (GENES-OF a7120)
  as length = (LENGTHS-OF LongSequences)
  when (length > 3000)
  Collect LongSequences)
```

or some variation of it. None worked although I'm sure I'm pretty close.

I also do not understand why I didn't get a list of genes when I used the "find-frames" command (function?), what exactly the value of this command?

From: Dr. E <[Virginia]>
To: DR. X <[Michigan]>
Date: Oct 21, 2004 7:53 AM
Subject: Re: Help with BioLingua

It's remarkable that you got as far as you have!
Here's one way to get a list of genes that you can then sift through by length:

```
(LOOP FOR frame IN (FIND-FRAMES "metabolism")
  AS genes = (GET-ELEMENT GO.related-genes FROM frame)
  WHEN (EXISTS genes)
  APPEND genes)
```

If you like what you get, you can save the result in a variable:

```
(ASSIGN metabolic-genes *)
```

The asterisk inserts the results of the previous operation). To find out how many genes you got:
...

From: Mr. M <[California]>
To: Dr. E <[Virginia]>
Date: Oct 21, 2004 9:34 AM
Subject: Re: Help with BioLingua

What are 'housekeeping' genes?

> (LOOP FOR frame IN (FIND-FRAMES "metabolism")
> AS genes = (GET-ELEMENT GO.related-genes FROM frame)
> WHEN (EXISTS genes)
> APPEND genes)

But this does not restrict the genes to the Anabaena 7120 organism.
You could do

APPEND (remove-if-not 'is-anabaena7120-gene genes)
  and
(defun is-anabaena7120-gene (gene) (eq ana7120 (#^Organism gene)))

==========

From: Dr. E <[Virginia]>
To: Mr. M <[California]>
Date: Oct 21, 2004 10:07 AM
Subject: Re: Help with BioLingua

>What are 'housekeeping' genes?

Housekeeping genes are those genes that are useful for the general maintenance of the cell under normal conditions. The term is usually used in the context "just houskeeping genes", implying "not interesting". But for those looking metabolism as a whole, they can be very interesting.

From: Mr. M <[California]>
To: DR. X <[Michigan]>
Date: Oct 22, 2004 11:13 AM
Subject: Re: Help with BioLingua

Here's an abbreviated script showing how to do exactly what you want, starting from after you found the GO.METABOLISM frame.

Hope this helps.

```
[...]
<2>> a7120
:: #$anabaena_pcc7120
<3>> (defun is-a7120-gene (g) (equal a7120 (#^Organism g)))
:: IS-A7120-GENE
<4>> (setq housekeeping-genes (#^Go.Related-Genes #$Go.Metabolism))
:: (#$A7120.alr7635 #$A7120.alr7622 #$A7120.all7592 #$A7120.alr7073
...)

<5>> (setq a7120-housekeeping-genes (remove-if-not 'is-a7120-gene
housekeeping-genes))
:: (#$A7120.alr7635 #$A7120.alr7622 #$A7120.all7592 #$A7120.alr7073
...)
<6>> (length housekeeping-genes)
:: 229
<8>> (setq result (loop for g in a7120-housekeeping-genes
when (> (length (extract-sequence g)) 3000)
collect g))
:: (#$A7120.alr3809 #$A7120.alr2680 #$A7120.alr2679 #$A7120.alr2678
#$A7120.all2649 #$A7120.all2648 #$A7120.all2647 #$A7120.all2646
#$A7120.all2645 #$A7120.all2644 #$A7120.all2643 #$A7120.all2642
#$A7120.all2635 #$A7120.all1695 #$A7120.all1649 #$A7120.all1648
#$A7120.all1643)
<9>> (length result)
:: 17
```

BioBike/KnowOS is a "Web 24.0" Platform:

"In developing BioBike, the biologists and computer scientists are developing a fundamental biological instrument—a biocomputational tool that must be used, and indeed is being used—by biologists to get real scientific work done—work that they could not get done any other way."

[A the same time they] are co-evolving a pidgin which exists [in both] their conversation, and [...] in the biocomputing platform [...].

The facility to dynamically extend the system's working vocabulary makes BioBike unique among computationally-based collaboration tools which, although they often support conversations among participants, do not usually themselves grow organically through these conversations.

Not merely learning to talk to one another, the scientists, engineers, and BioBike are doing real work of biocomputation and at the same time as they are evolving the way that this work gets done, they are extending their own understandings, amoeba-like into one another's areas of expertise.

Specialized programming platforms are becoming increasingly important as computers infuse greater parts of our daily lives, and as we wish to have greater control over them. [...] the programming languages that are the heart of computing platforms serve as, at the same time, inter-languages in the trading zones that are these platforms, and that the functions and objects of those languages serve as boundary objects in these trading zones. [...] the participants in the collaboration co-evolve the BioBike inter-languages themselves..."

J Shrager, in press, The Evolution of BioBike: Community Adaptation of a Biocomputing Platform; Studied in the History and Philosophy of Science.

BioBike/KnowOS Integrates Scientists and Computation in a *Trading Zone*

Simulation: What does this model predict?

Explanation: How does it make these predictions?

Model Identification: What models fit this data?

Collaboration: Interact with scientists...
...and help scientists interact with one another!

-- Inference sharing and peer group critical analysis

-- Ability to track the chain of inference

Trading Zones and the
Bayes Community Model

New Knowledge

Client/server architecture permits collaboration among scientists through "publication" of hypothesis and linking in as evidence

Incoming knowledge are distributed to the scientists according to the hypotheses they are working, and heuristically knitted into the ongoing model development process

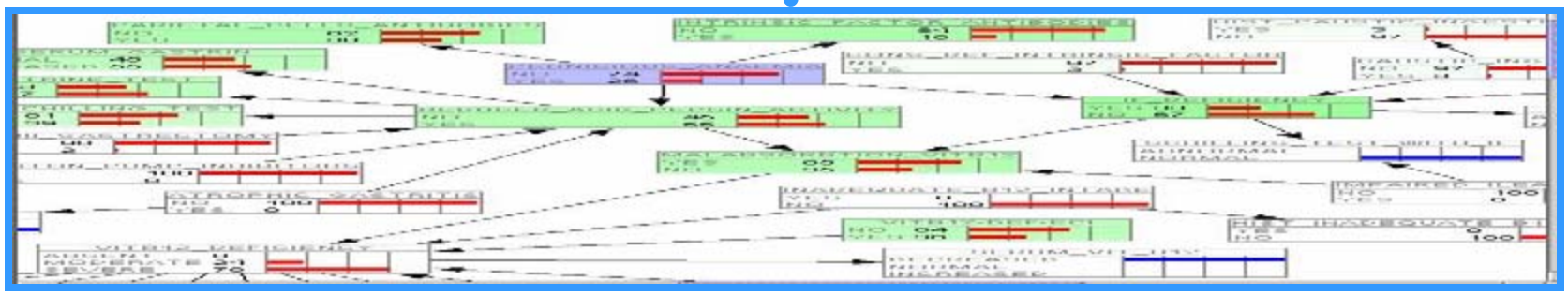
Analysis of Competing Hypotheses

Analysis of Competing Hypotheses

Analysis of Competing Hypotheses

Analysis of Competing Hypotheses

Linked matrices project a Bayesian influence network



ACH:

Analysis of Competing Hypotheses

Sort Evidence By: Order Added | Create Matrix | Show Tutorial

		Code	Type	Weight	H: 1	H: 2	H: 3	H: 4	H: 5		
					Disgruntled Michael's employee or customer	Foreign terrorist(s)	Two black males in blue car with unknown, possibly criminal, motives	Lone serial killer (almost certainly male, 80% prob white	Domestic terrorist(s), white militiamen		
	Inconsistency Score				-89	-55	-5	-47	-43		
	Create Evidence										
E43	22 Oct Johnson shot on steps of Ride-on-Bus		Police report	HIGH	I	C	C	C	C		
E42	Accent sounded Caribbean, Jamaican?		From sniper(s)	LOW	N	I	I	I	I		
E41	Sniper calls from Ponderosa; five red stars note		From sniper(s)	HIGH	I	I	C	I I	I		
E40	Ala police report Armalite catalog dropped by suspect		Police report	HIGH	NA	NA	C	NA	NA		
E39	Accent of phone caller Hispanic or Jamaican		Analysis	MEDIUM	N	I	I	I	I		
E38	Sniper calls Pastor Sullivan (2 men, accent, Ala)		From sniper(s)	MEDIUM	I	I	C	C	I		
E37	Credit card used in Tacoma WA linked with Alabama		Police report	MEDIUM	I	I	C	I	I		
E36	Cinnaraison snack bag		Analysis	LOW	NA	NA	NA	NA	NA		
E35	Handwriting matches Tarot card		Analysis	HIGH	I	I	C	C	I		
E34	Ziploc bag with letter (use of "we") and demands		From sniper(s)	HIGH	I I	I I	C C	I I	I I		
E33	.233 casing found consistent in most cases		Analysis	LOW	C	C	C	C	C		
E32	19 Oct Hopper shot at Ashlawn Ponderosa		Police report	HIGH	I	C	C	C	C		
E31	Sniper calls Baliles; provides Alabama info (2 suspects)		From sniper(s)	HIGH	I I	I	C C	I I	I		
	Sniper calls dispatcher		From sniper(s)								

Trading Zones and the
Bayes Community Model

Scientists can “promote” hypotheses as if they were results, and other scientists can import these.
The system automatically tracks provenance (code+params, or BioDeducta “explanations”) to build a network of support.

user: Shrager:

Top Matrix	Weight:	3301930422: Foreign Terrorist	33 Wo Sn
Operations:			
3301930131: Terrorists sniping	MEDIUM	CC	
3301917159: Accent sounded Caribbean, Jamaican?	MEDIUM	C	
3301917788: TV profile: white, male, military background	MEDIUM	I	
Support:		0.5	-1.

user: Heuer:

Top Matrix	Weight:	3301930556: Terrorists sniping	3301929911: Kids out joyriding with guns	Diagnosticity:
Operations:		Delete / Promote		
3301917397: Forensics shows all shot at long range with .223 bullets	MEDIUM		II	1 confirm, and 1 disconfirm (0)
3301917333: 2 Oct shot fired thru Michael's store window	MEDIUM	I	CC	1 confirm, and 1 disconfirm (0)
Support:		0.5	0.0	

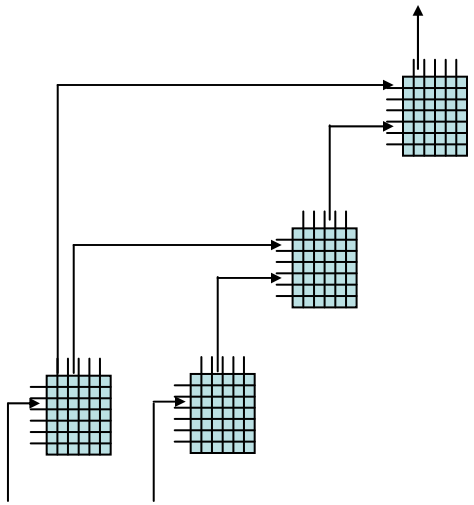
When the support for linked results changes, results that depend upon those are likewise changed in level of belief, or are flagged for reconsideration.

user: Shrager:

Top Matrix	Weight:	3301930422: Foreign Terrorist	33 Wo Sn
Operations:			
3301930131: Terrorists sniping	MEDIUM	CC	
3301917159: Accent sounded Caribbean, Jamaican?	MEDIUM	C	
3301917788: TV profile: white, male, military background	MEDIUM	I	
Support:		0.5	-1.

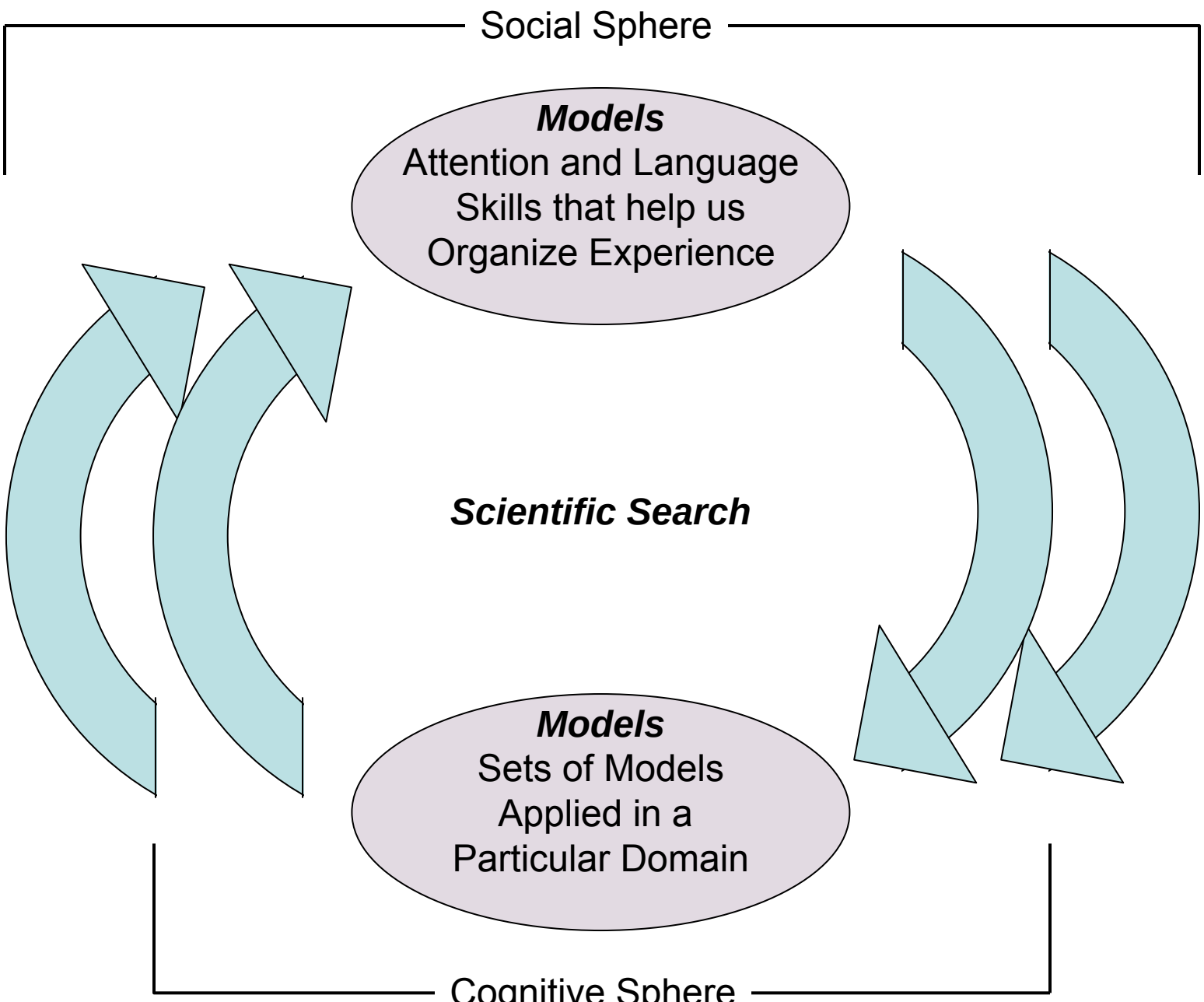
user: Heuer:

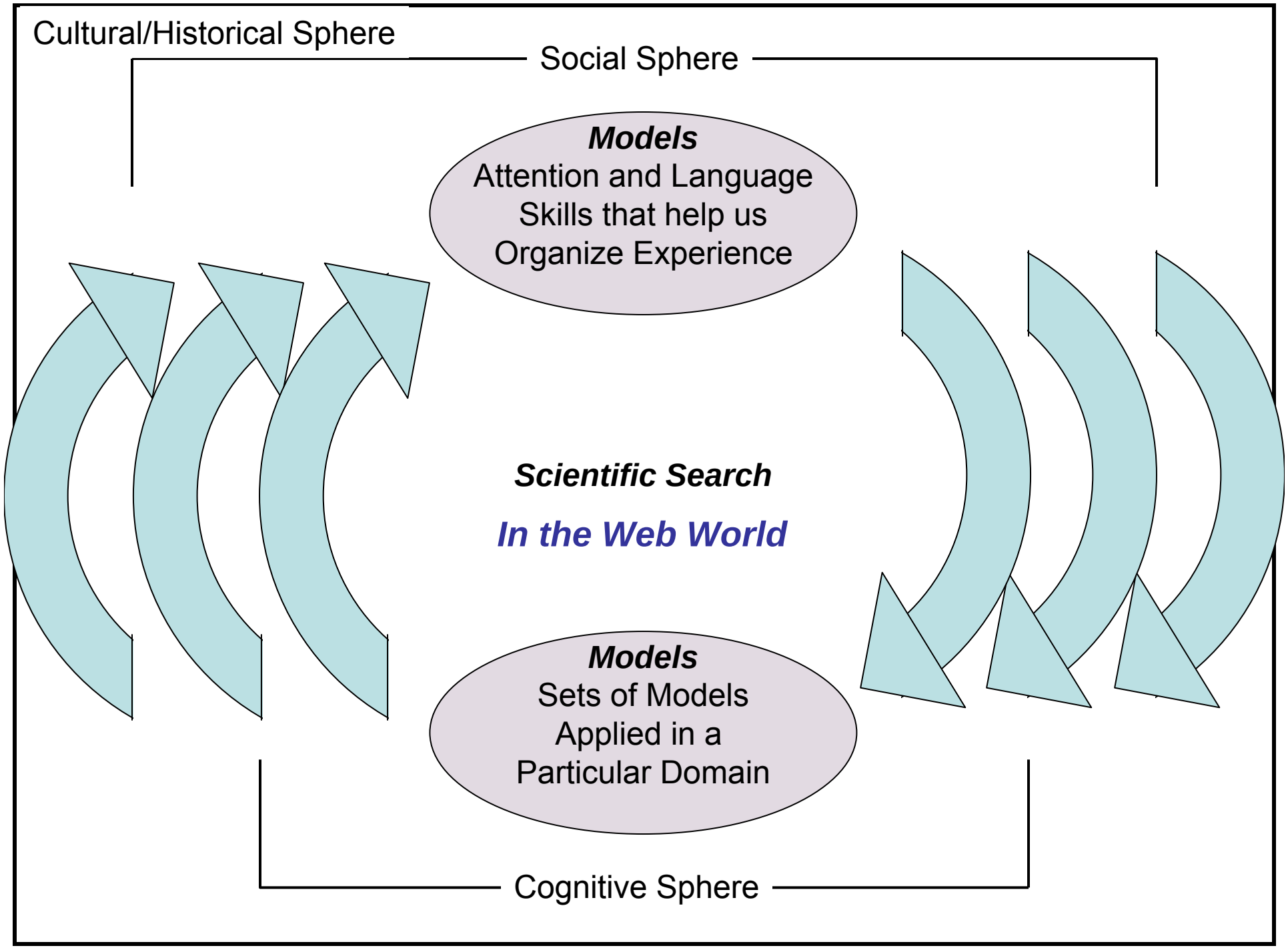
Top Matrix	Weight:	3301930556: Terrorists sniping	3301929911: Kids out joyriding with guns	Diagnosticity:
Operations:		Delete / Promote		
3301917397: Forensics shows all shot at long range with .223 bullets	MEDIUM		II	1 confirm, and 1 disconfirm (0)
3301917333: 2 Oct shot fired thru Michael's store window	MEDIUM	I	CC	1 confirm, and 1 disconfirm (0)
Support:		0.5	0.0	



-- Inference sharing and peer group critical analysis

-- Ability to track the chain of inference


Social Sphere
Models
Attention and Language
Skills that help us
Organize Experience
Scientific Search
Models
Sets of Models
Applied in a
Particular Domain
Cognitive Sphere


Cultural/Historical Sphere
Social Sphere
Models
Attention and Language
Skills that help us
Organize Experience
Scientific Search
In the Web World
Models
Sets of Models
Applied in a
Particular Domain
Cognitive Sphere

How Science Thinks:
The Science and Engineering of Science and Engineering

BioBike/KnowOS:
JP Massar
Andrew Pohorille
Mike Travers
Jeff Elhai
Richard Waldinger

Afferent:
David Chapman
David Gladstein
Randy Gobbel
Jon Handler
Mike Travers

Cyclodyn Experiments:
Kevin Arrigo
Stephen Bay
Devaki Bhaya
Arthur Grossman
Rochelle Labiosa
Tasha Reddy
CJ Tu

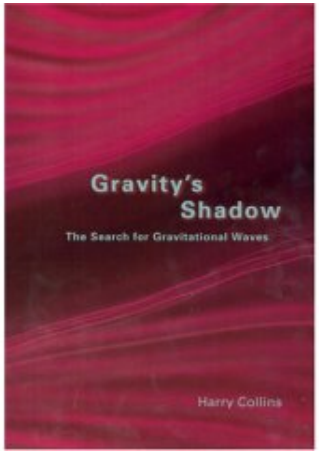
CACHE:
JP Massar
Peter Pirolli
Dorrit Billman
Gregorio Convertino

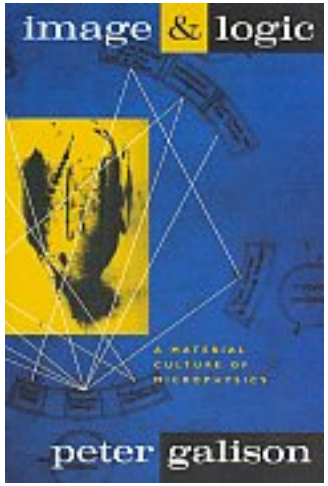
BioDiscovery:
Stephen Bay
Lonnie Chrisman
Pat Langley
Andrew Pohorille
Kazumi Saito
Richard Waldinger

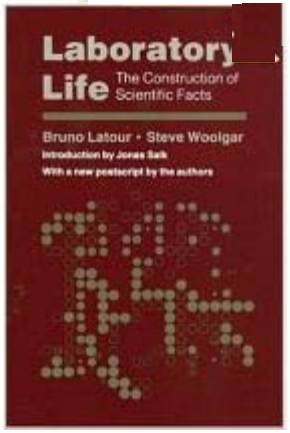
Funding from NASA, NSF, Carnegie Inst. DPB, Franz Inc., Lispworks Inc. and others.

How Science Thinks:

The Science and Engineering of Science and Engineering







How Science Thinks:

The Science and Engineering of Science and Engineering