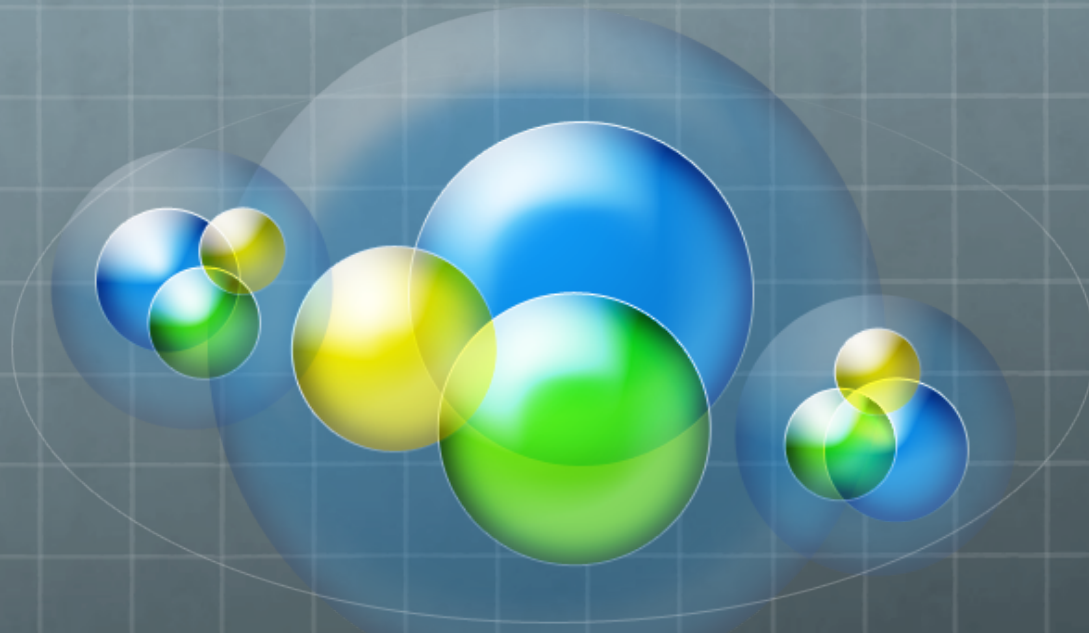


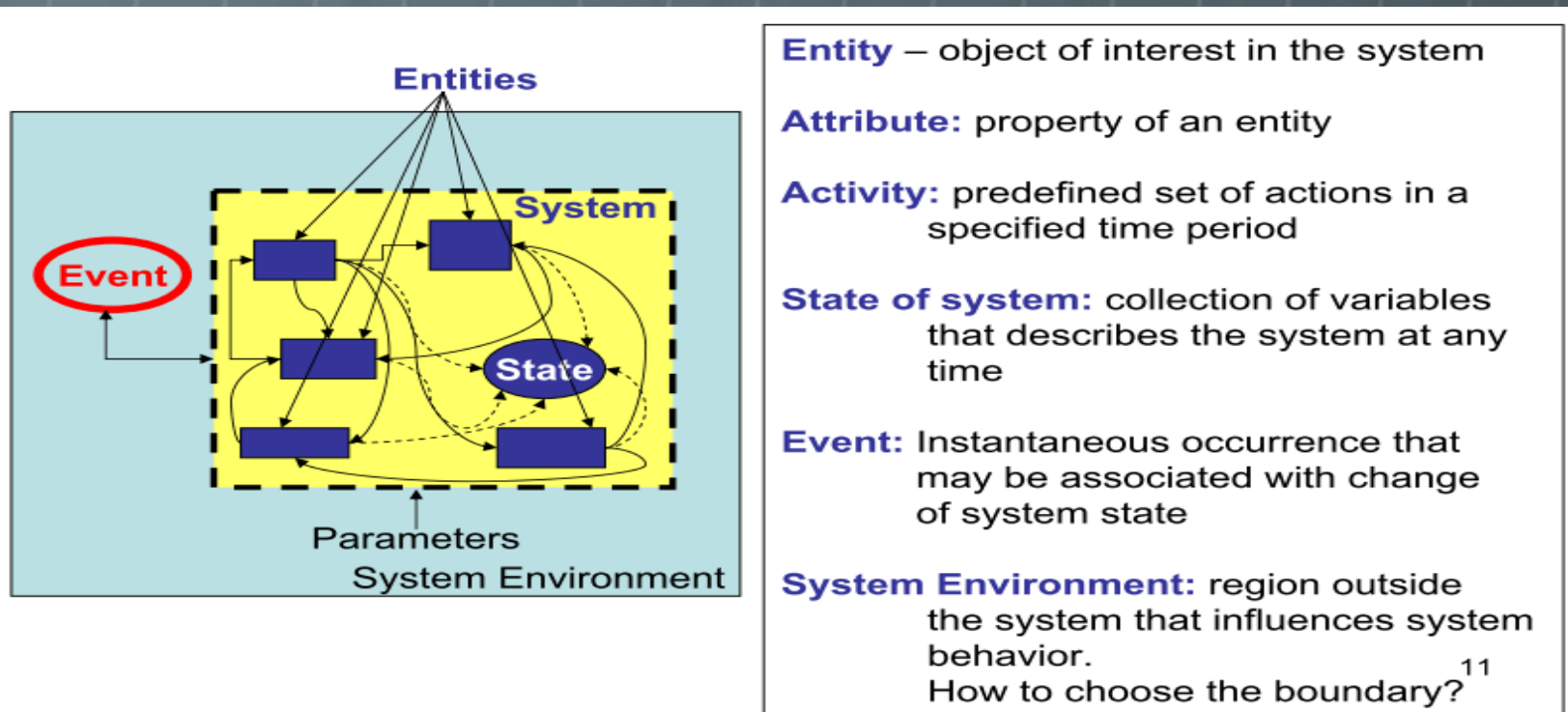


Knowledge Discovery in System Biology

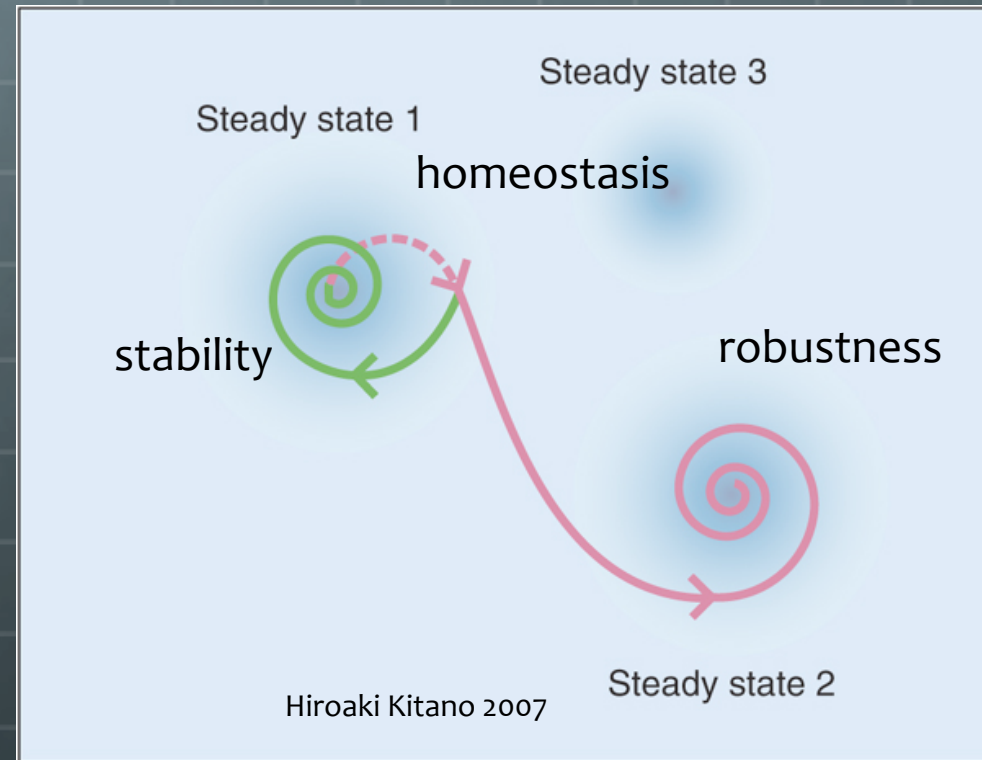
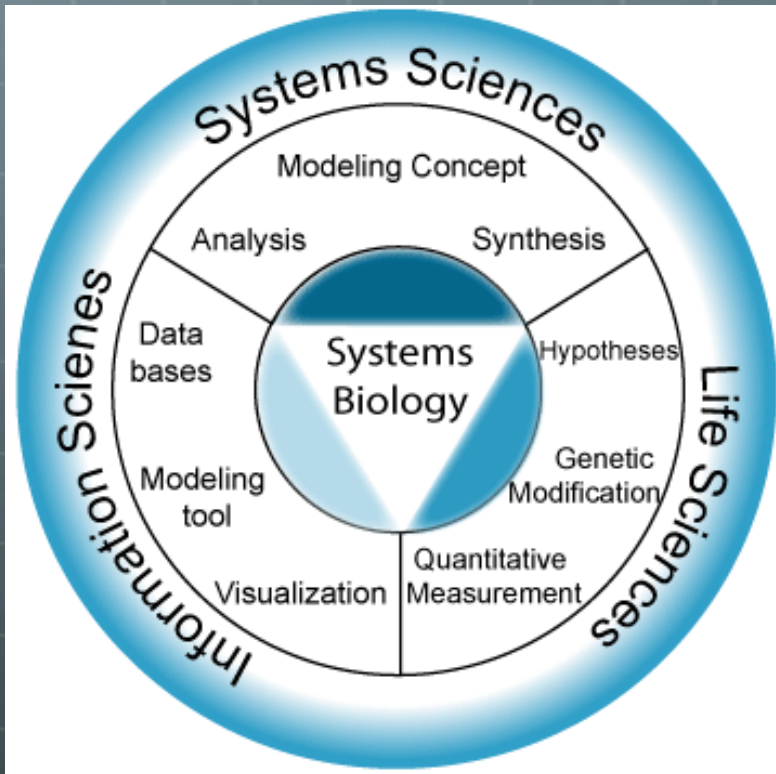
Katsumi Inoue
Yoshitaka Yamamoto
Andrei Doncescu

System, System Biology, Complex Systems

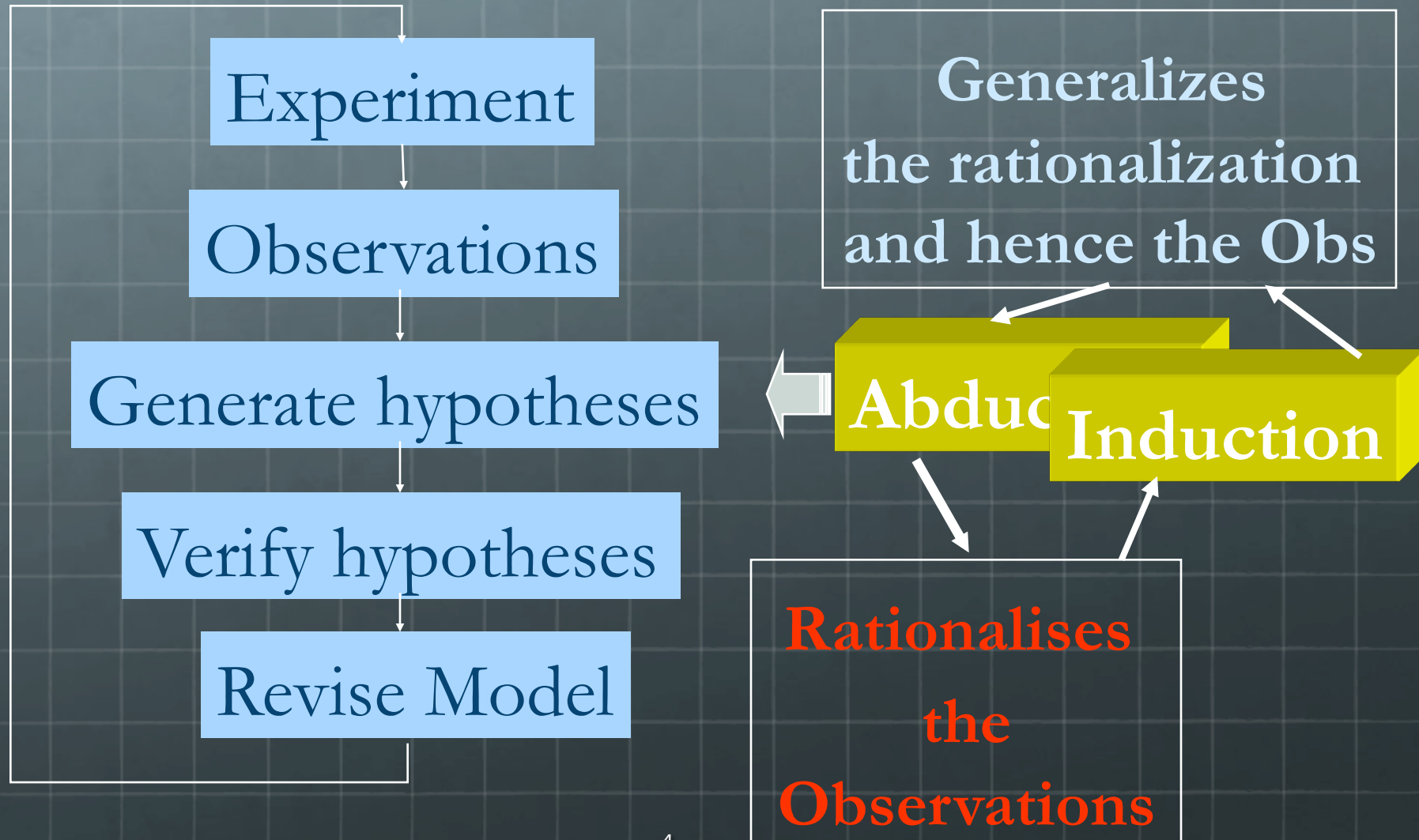
- 🌐 **System:** set of objects, joined to accomplish some purpose



Stability, Homeostasis and Robustness

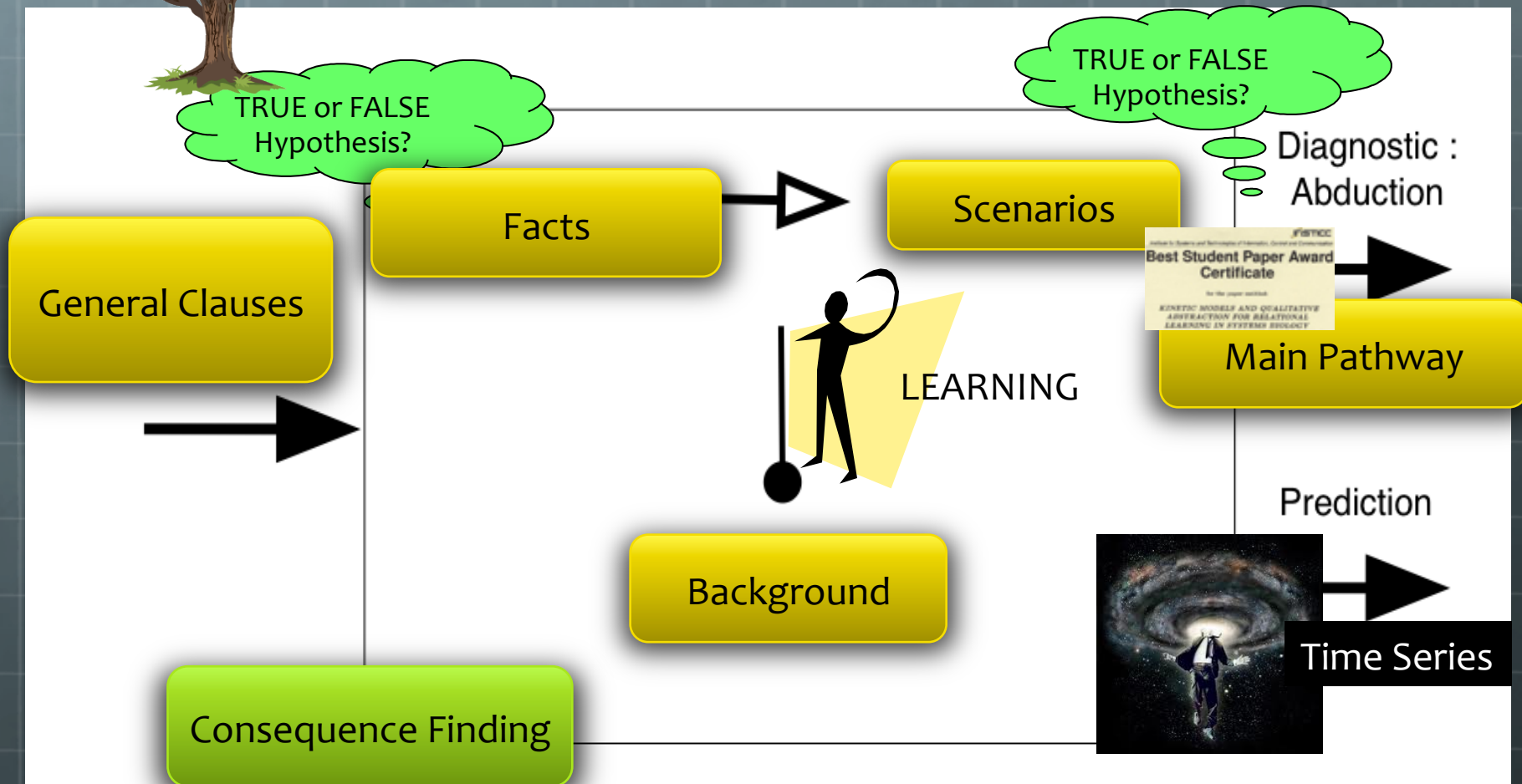


Logical Reasoning for Scientific Analysis

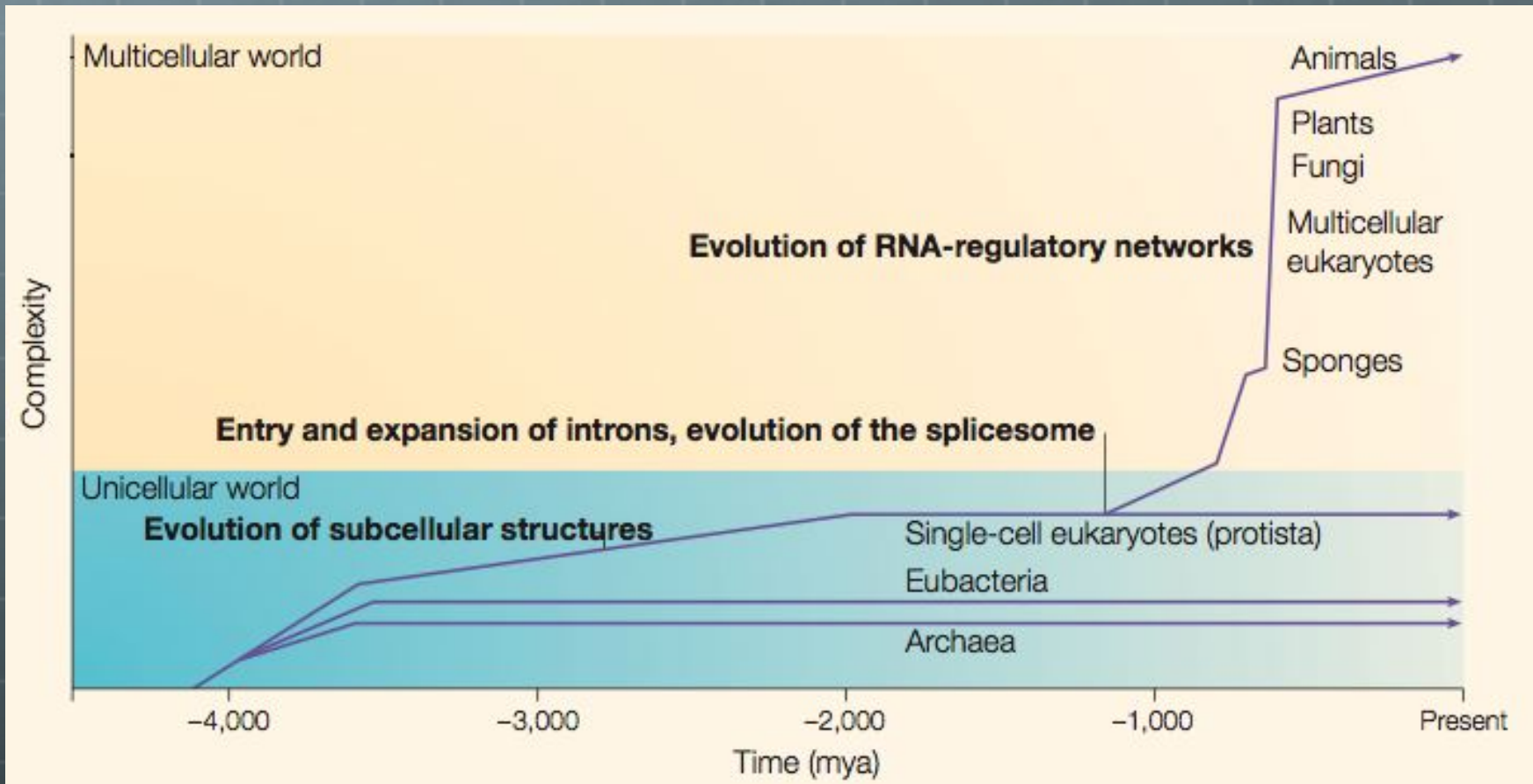


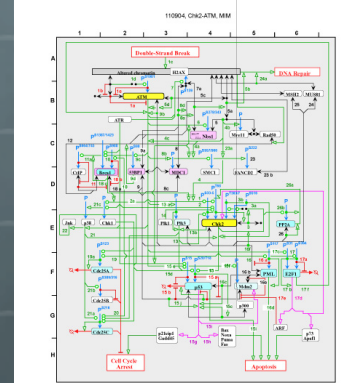
Knowledge-Based Discovery :

Ability to find non-trivial consequences of an axiom set



A Simplified History of Life on Earth

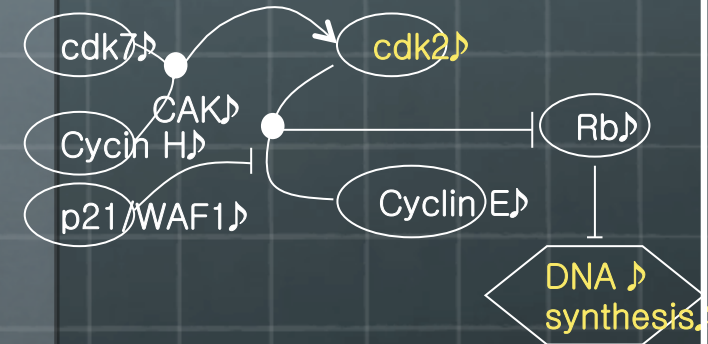




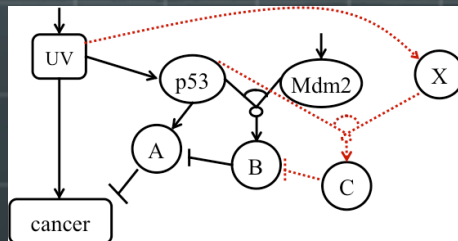
Observations

Hypotheses

$B \wedge H \models$



$$E_{i+1} = Th(E_i) \cup \left\{ B / \left(\frac{A : B}{B} \right) \in D, A \in E_i, \neg B \notin E_i \right\}$$



Traditional science

Automated science

Hypotheses

Machine-encoded logical hypotheses

Chemical knowledge

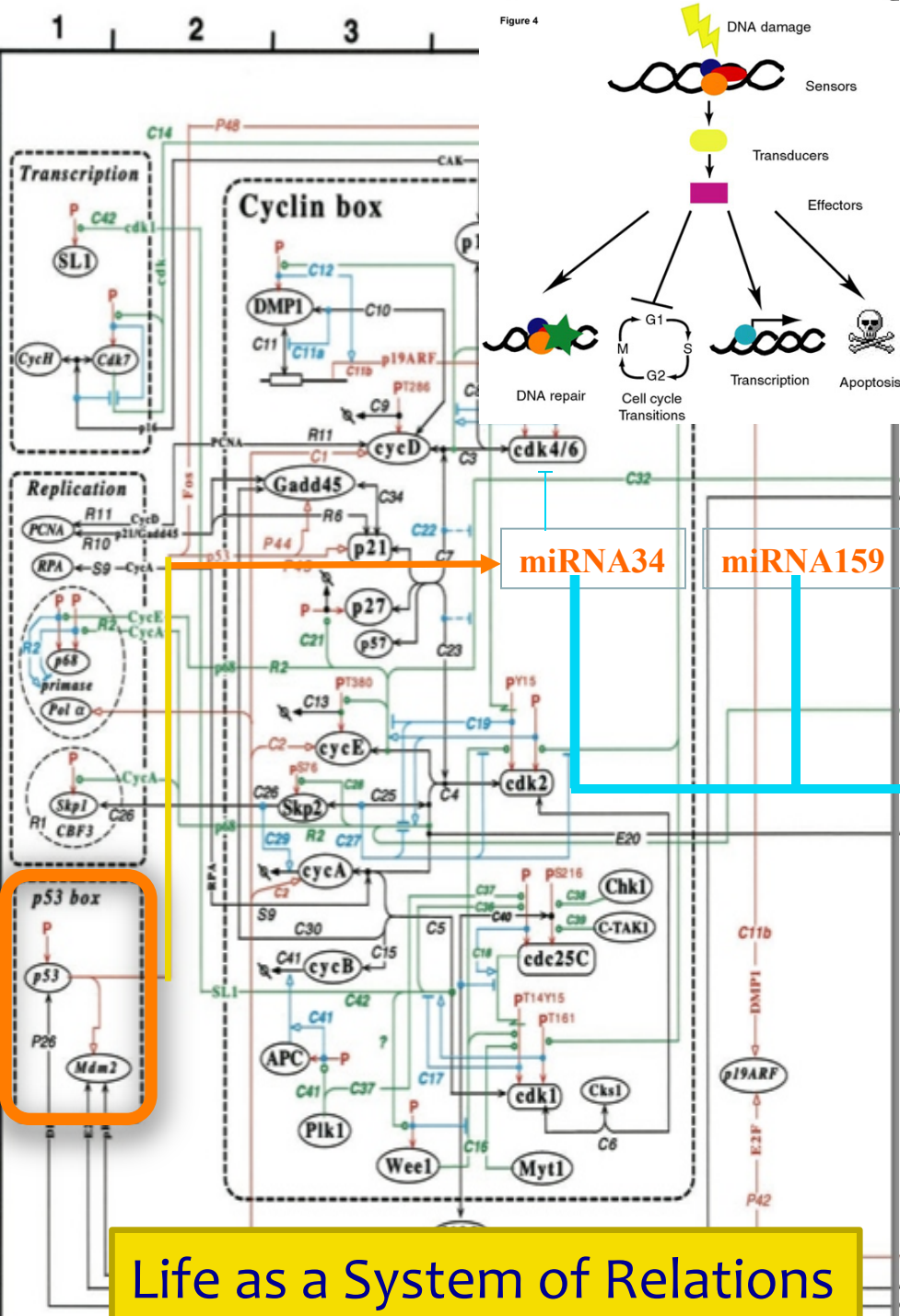
Machine-encoded chemical algebra

Experiments

Chemical Turing machine programs

Experimental design

Decision theory



Molecular Interaction Map of the Mammalian Cell Cycle Control and DNA Repair Systems

Kurt W. Kohn*

E2F box

Myc box

Chromatin & acetylase box

```
[+product(p_atm_bound), +phosphorylation(p_atm_bound,nbs1)]
[+product(p_nbs1), -phosphorylation(p_atm_bound,nbs1)]
[+product(p_fancd2), -phosphorylation(p_atm_bound,fancd2)]
[+product(p_fancd2), +stimulation(p_fancd2,cell)]
[-product(cell_cycle_arrest), -stimulation(p_fancd2,cell)]
[+binding(chk2_53bp1,mus81), -product(chk2_53bp1)]
[+product(mus81_chk2), -binding(chk2_53bp1,mus81)]
[-product(mus81_chk2), +binding(mus81_chk2,altered_dna)]
[+product(mus81_bound_to_dna), -binding(mus81_chk2,altered_dna)]
[+ans(_0_1), +product(altered_dna), -phosphorylation(_0_1)]
```

Top clauses:

```
[+ans(_0_1), +product(altered_dna), -phosphorylation(_0_1)]
Production field:
```

```
[+phosphorylation(_0_1), +product(altered_dna), +ans(_0_1)]
```

```
Stage 1 (dep:1 len:- inf:0)
Stage 2 (dep:2 len:- inf:8)
Stage 3 (dep:3 len:- inf:85)
Stage 4 (dep:4 len:- inf:295)
Stage 5 (dep:5 len:- inf:596)
Stage 6 (dep:6 len:- inf:1080)
Stage 7 (dep:7 len:- inf:1702)
Stage 8 (dep:8 len:- inf:2785)
Stage 9 (dep:9 len:- inf:4053)
```

MicroARN Analysis for Cancer Relapse

A **logic-based machine learning** technique

- Richer representation formalisms (First-order predicate logic)
- Classification

■ Input:

- B : background theory.
- E : observations.
- LB : language bias for restricting the syntax of hypotheses

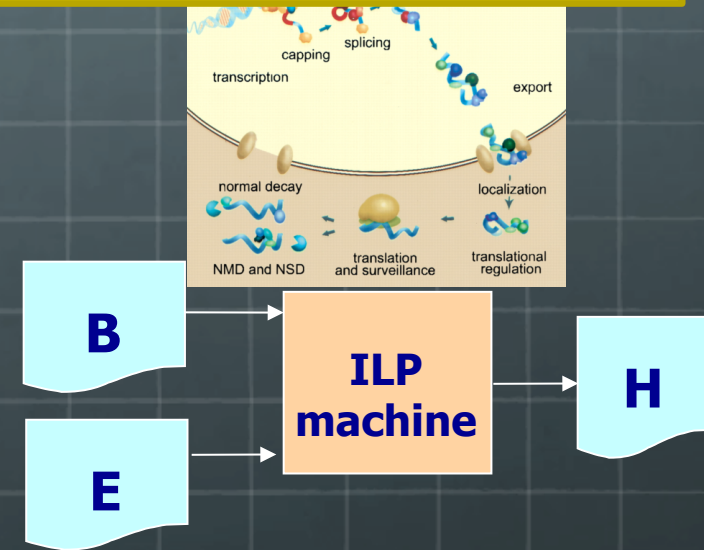
■ Output:

- H : hypothesis satisfying that
 - H is a clause belonging to LB
 - $B \wedge H$ logically explains E

Background

$\text{Breslow}(X, \text{high}) \wedge \text{mir21}(X, \text{high}) \wedge \text{miR-222}, \text{miR-23a}(X, \text{high}) \wedge \text{miR-92}(X, \text{high})$
 $\wedge \text{miR-149}(X, \text{high}) \wedge \text{miR-221}(X, \text{high}) \rightarrow \text{relapse}(X)$

Life as a System of Relations



Results using Inoue's ILP system

- Input file: the information on 15 patients with melanoma / 5 healthy patients
- Causal relations between status of relapsing and 19 factors

Hypothesis (Clause)	Rs (%)	Rf (%)	Expert's opinion
(7) [mir182(X, low), metastasis(X), age(X, medium) → relapse(X)]	50	8	○
(6) [metastasis(X), age(X, medium) → relapse(X)]	40	10	○
(7) [mir630(X, low), mir182(X, high) → relapse(X)]	30	10	

Thank you

