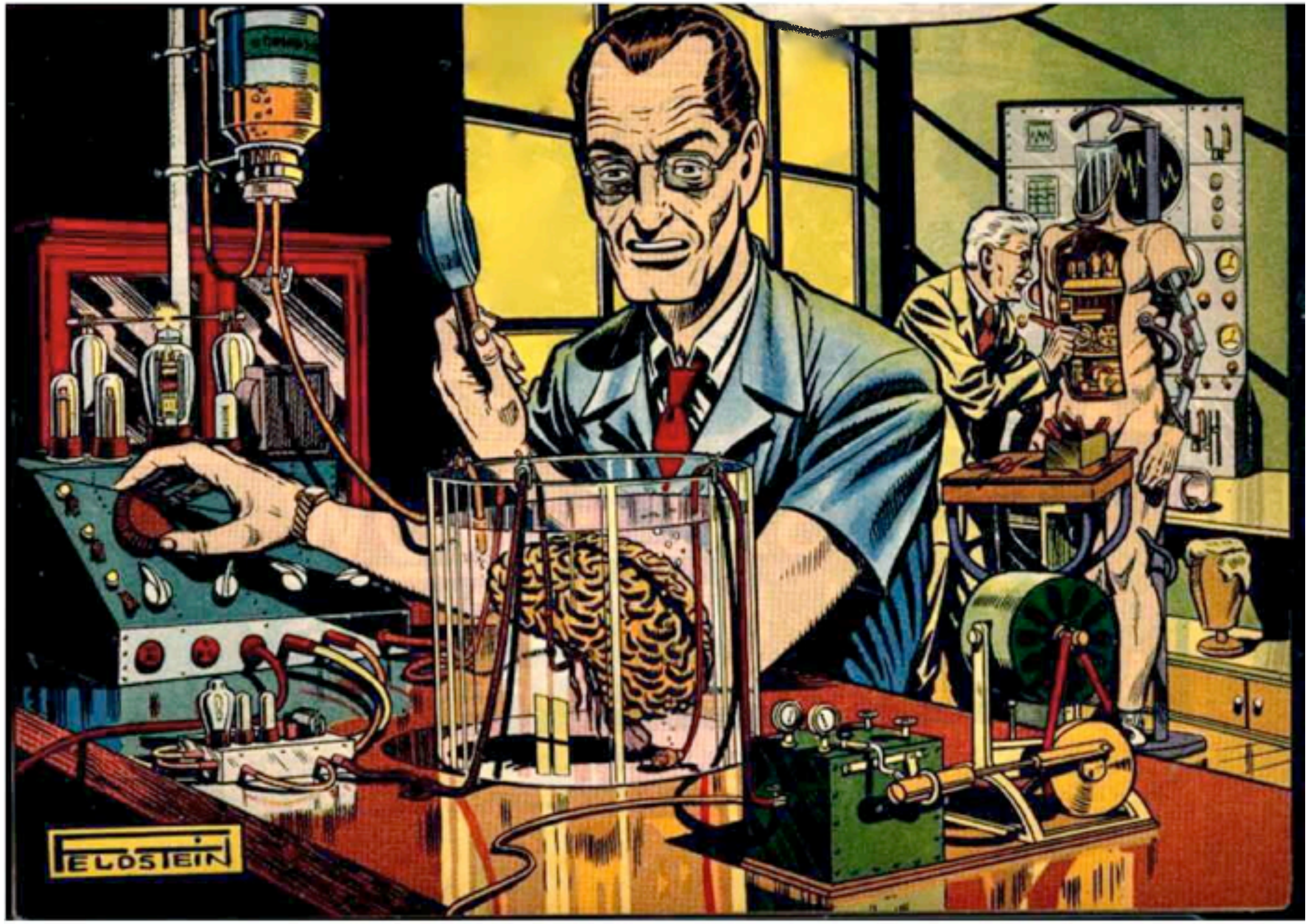
Active Sensing and Learning



ICASSP 2011, May 23, Prague

www.ece.wisc.edu/~nowak/ASL.html

Jarvis Haupt
www.ece.umn.edu/~jdhaupt

Rob Nowak
www.ece.wisc.edu/~nowak

Adaptive Information

Goal: Estimate an unknown object $x \in \mathcal{X}$ from scalar samples

Information: samples of the form $y_1(x), \ldots, y_n(x)$,
the values of certain functionals of x

Adaptive Information

Goal: Estimate an unknown object $x \in \mathcal{X}$ from scalar samples

Information: samples of the form $y_1(x), \ldots, y_n(x)$, the values of certain functionals of x

Non-Adaptive Information: $y_1, y_2, \cdots \in \mathcal{Y}$ non-adaptively chosen (deterministically or randomly) independent of x

Adaptive Information

Goal: Estimate an unknown object $x \in \mathcal{X}$ from scalar samples

Information: samples of the form $y_1(x), \ldots, y_n(x)$, the values of certain functionals of x

Non-Adaptive Information: $y_1, y_2, \cdots \in \mathcal{Y}$ non-adaptively chosen (deterministically or randomly) independent of x

Adaptive Information: $y_1, y_2, \cdots \in \mathcal{Y}$ are selected sequentially and y_i can depend on previously gathered information, i.e., $y_1(x), \ldots, y_{i-1}(x)$

Adaptive Information

Goal: Estimate an unknown object $x \in \mathcal{X}$ from scalar samples

Information: samples of the form $y_1(x), \dots, y_n(x)$,
the values of certain functionals of x

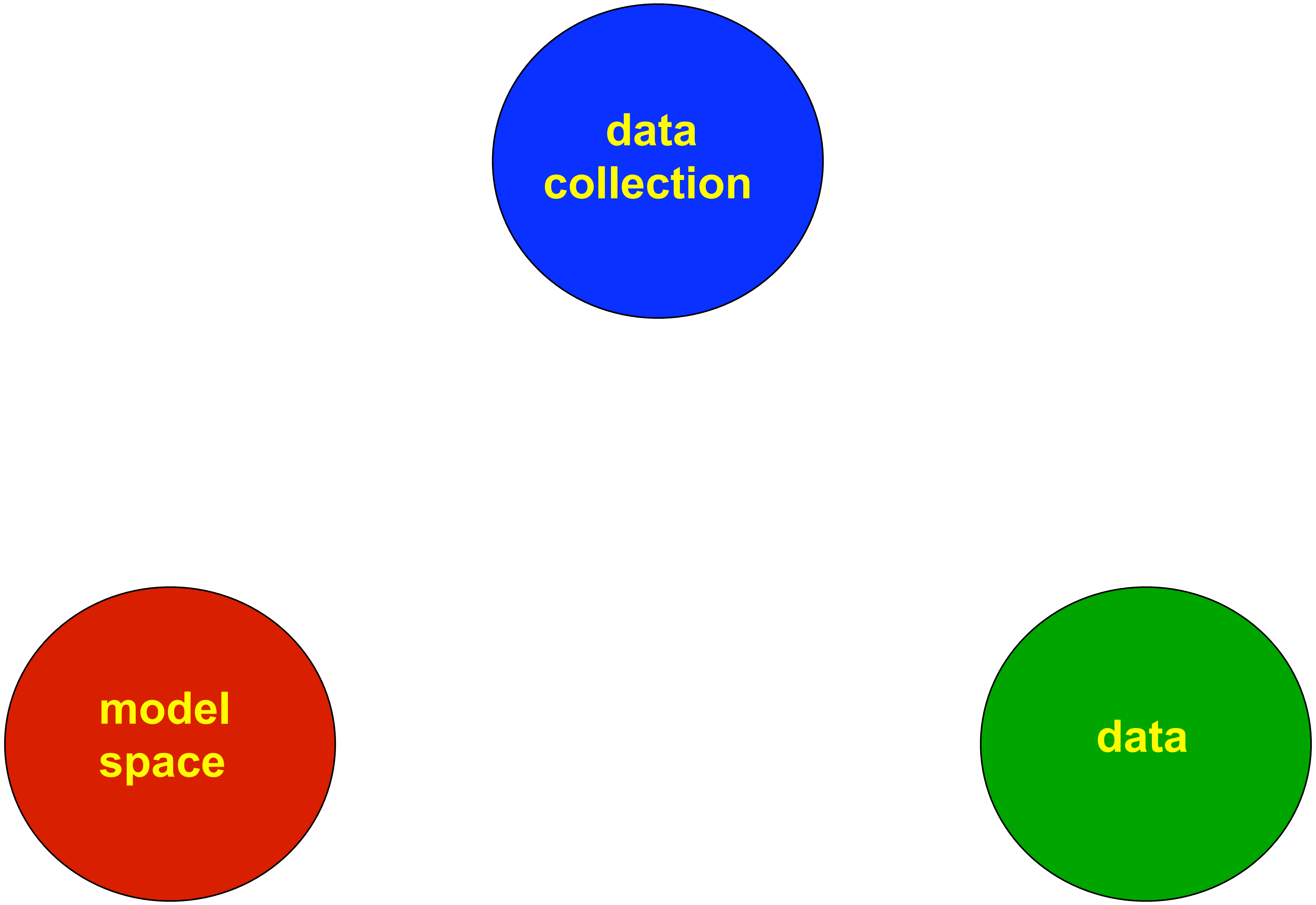
Non-Adaptive Information: $y_1, y_2, \dots \in \mathcal{Y}$ non-adaptively chosen (deterministically or randomly) independent of x

Adaptive Information: $y_1, y_2, \dots \in \mathcal{Y}$ are selected sequentially and y_i can depend on previously gathered information, i.e., $y_1(x), \dots, y_{i-1}(x)$

Does adaptivity help?

Feedback from Data Analysis to Data Collection

$\mathcal{Y}$: possible measurements/experiments

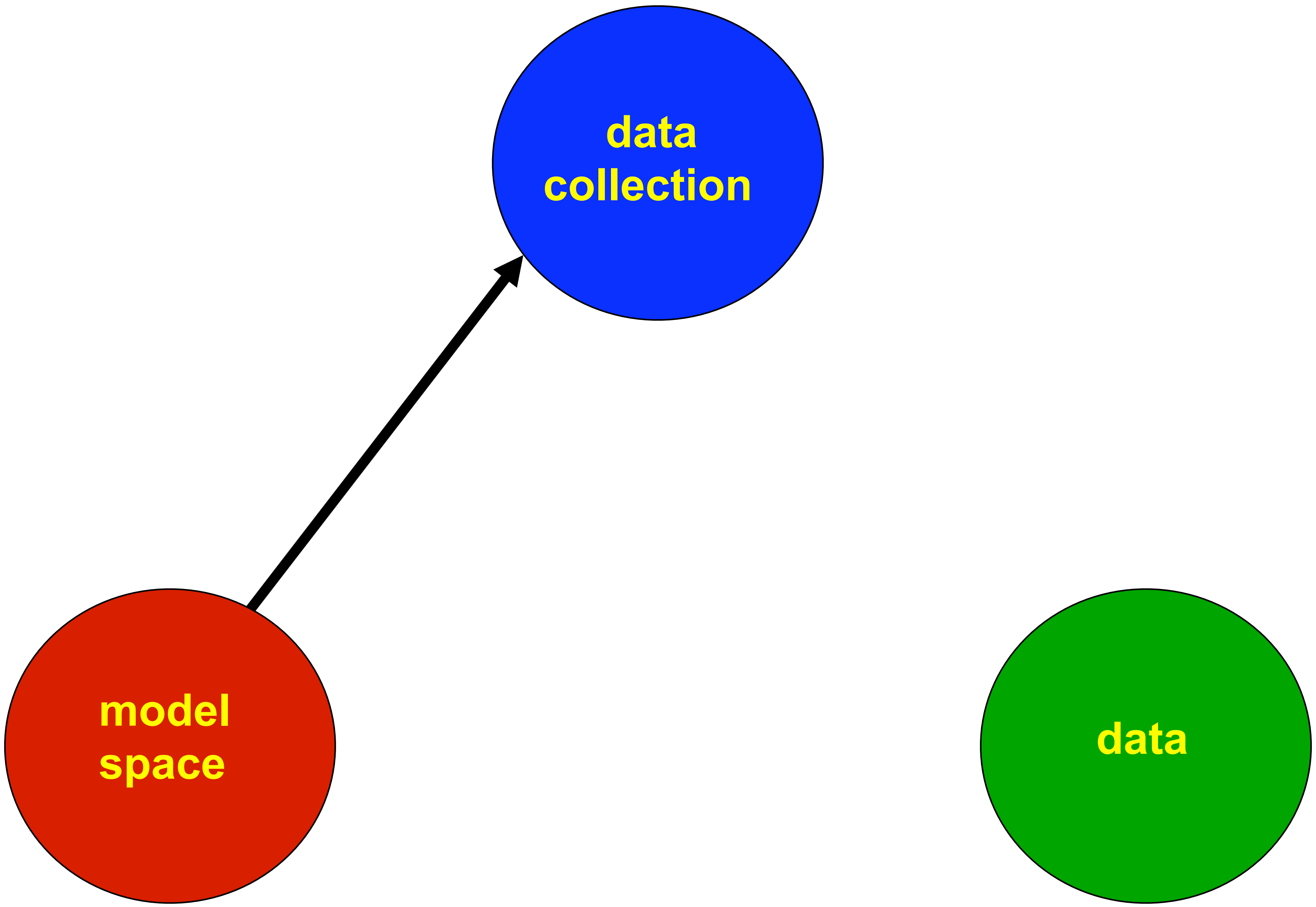


$\mathcal{X}$: models/hypotheses
under consideration

$y_1(x), y_2(x), \ldots$: information/data

Feedback from Data Analysis to Data Collection

$\mathcal{Y}$: possible measurements/experiments

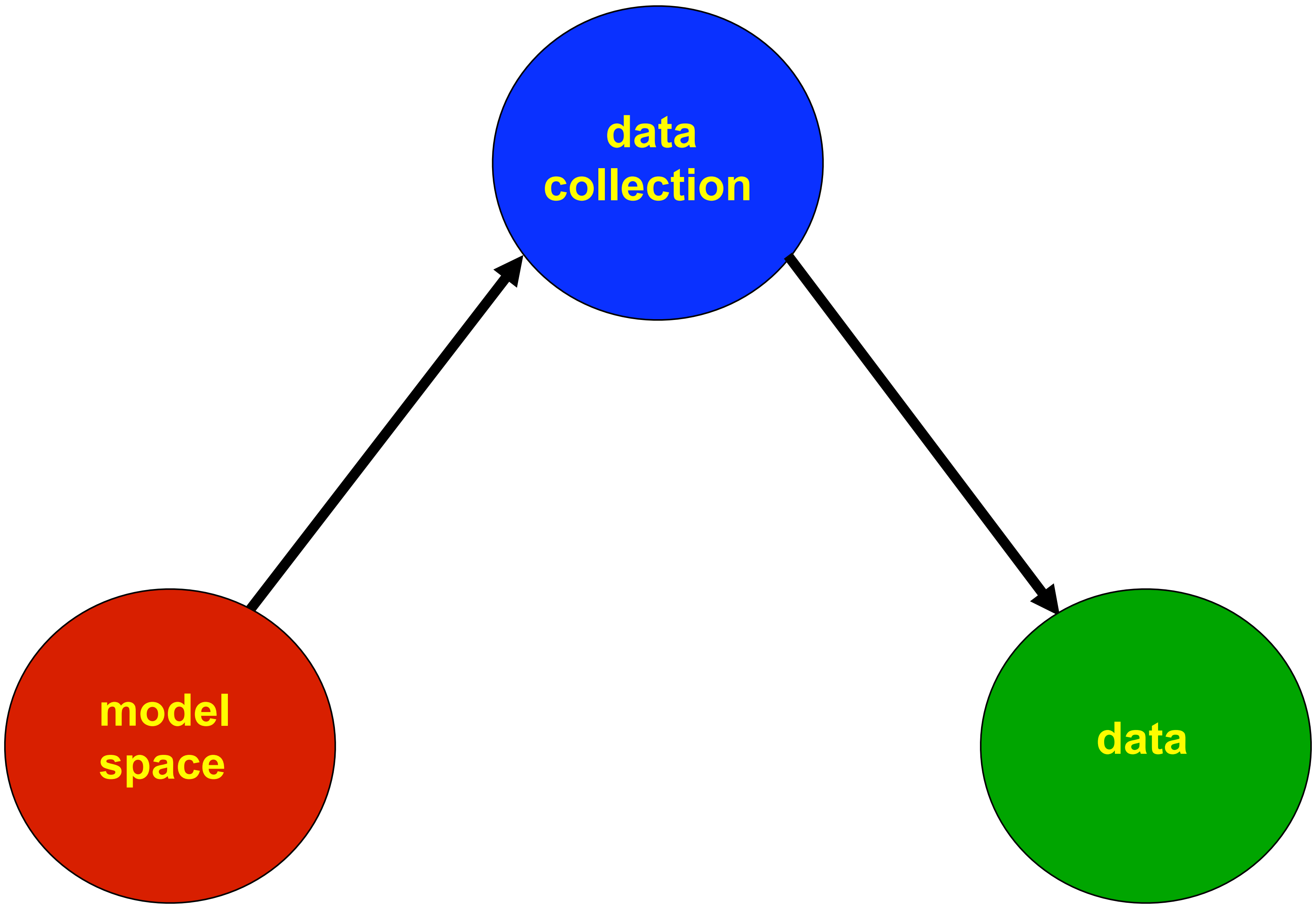


$\mathcal{X}$: models/hypotheses under consideration

$y_1(x), y_2(x), \ldots$: information/data

Feedback from Data Analysis to Data Collection

$\mathcal{Y}$: possible measurements/experiments

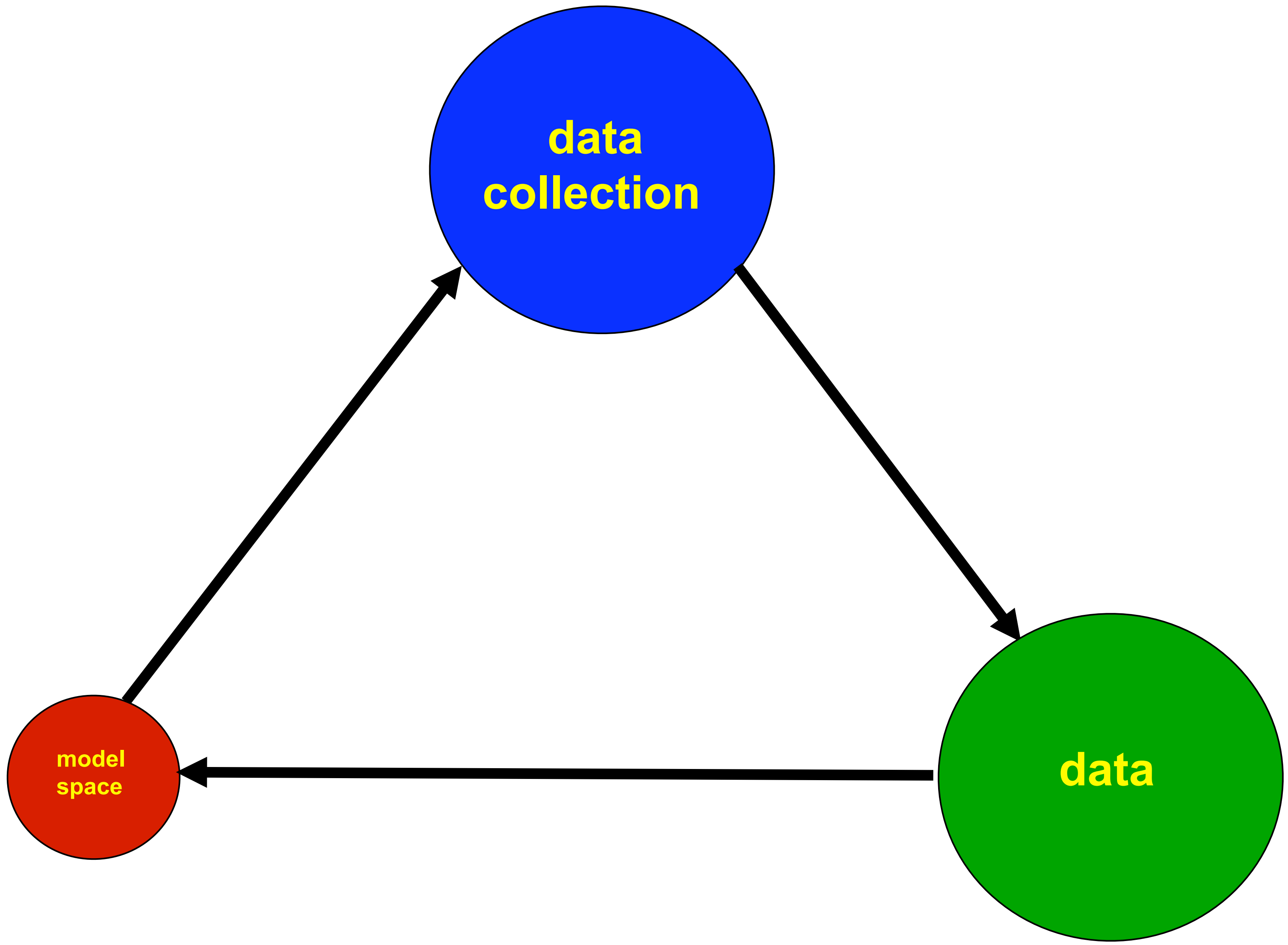


$\mathcal{X}$: models/hypotheses under consideration

$y_1(x), y_2(x), \ldots$: information/data

Feedback from Data Analysis to Data Collection

$\mathcal{Y}$: possible measurements/experiments



$\mathcal{X}$: models/hypotheses
under consideration

$y_1(x), y_2(x), \ldots$: information/data

Outline of Tutorial

Part 1: Introduction (Rob), 9:00-9:30

Part 2: Active Sensing (Jarvis), 9:30-10:30

Break, 10:30-10:45

Part 3: Active Learning (Rob), 10:45-11:45

Part 4: Conclusions and Future Directions, 11:45-12

Outline of Part 1:

Sequential Experimental Design

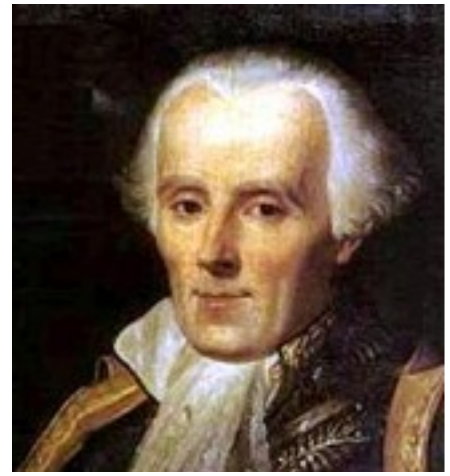
Adaptive Sensing for Sparse Recovery

Sensing and Inference in Large Networked Systems

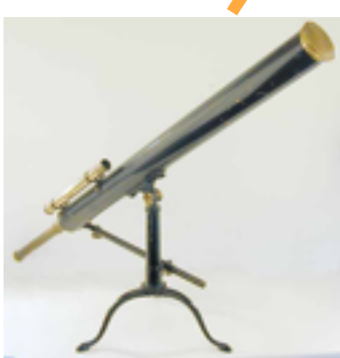
Active Learning in Machines and Humans

Mathematics of Active Sensing and Learning

Sequential Experimental Design



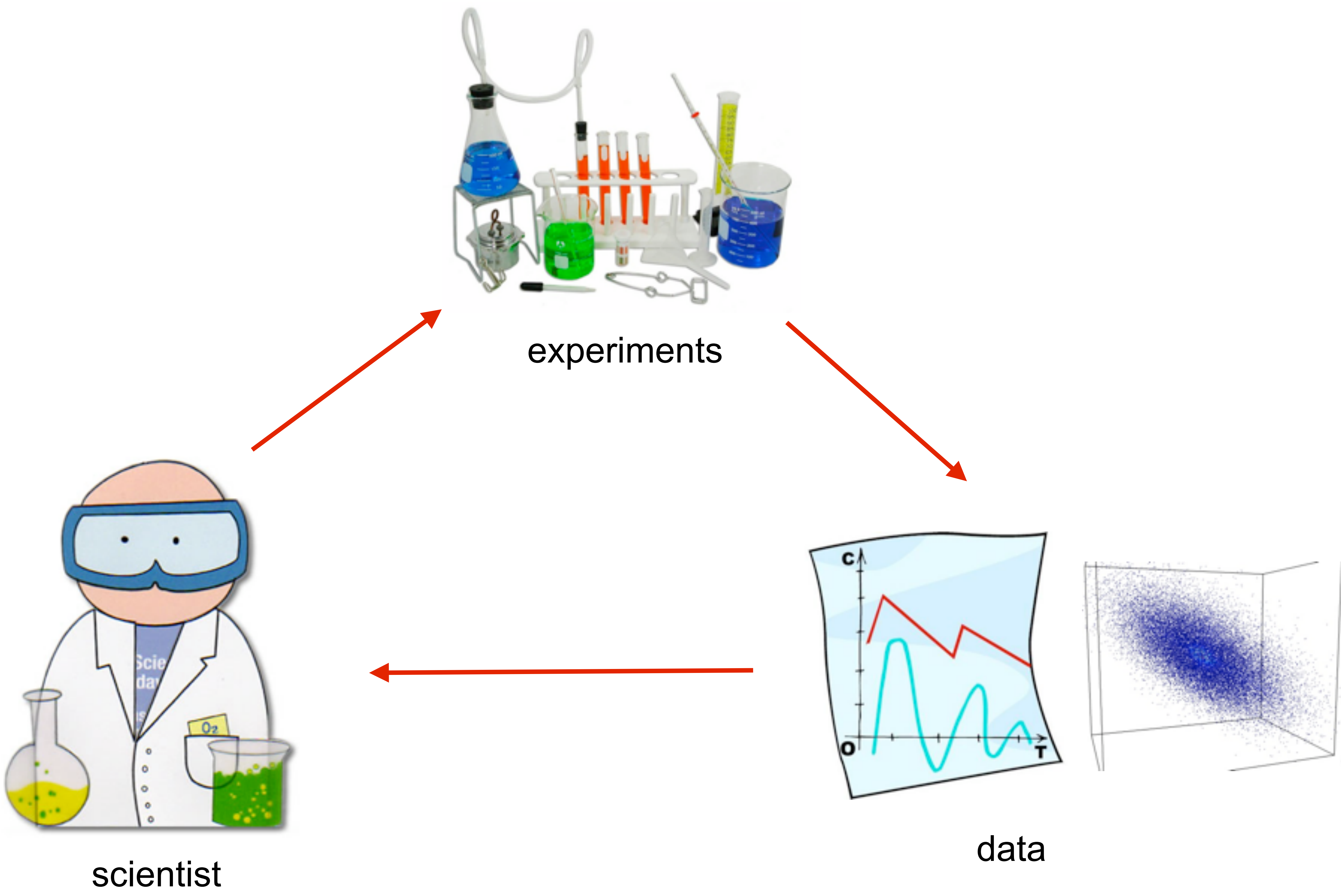
Decided to make new astronomical measurements when “the discrepancy between prediction and observation [was] large enough to give a high probability that there is something new to be found.” Jaynes (1986)



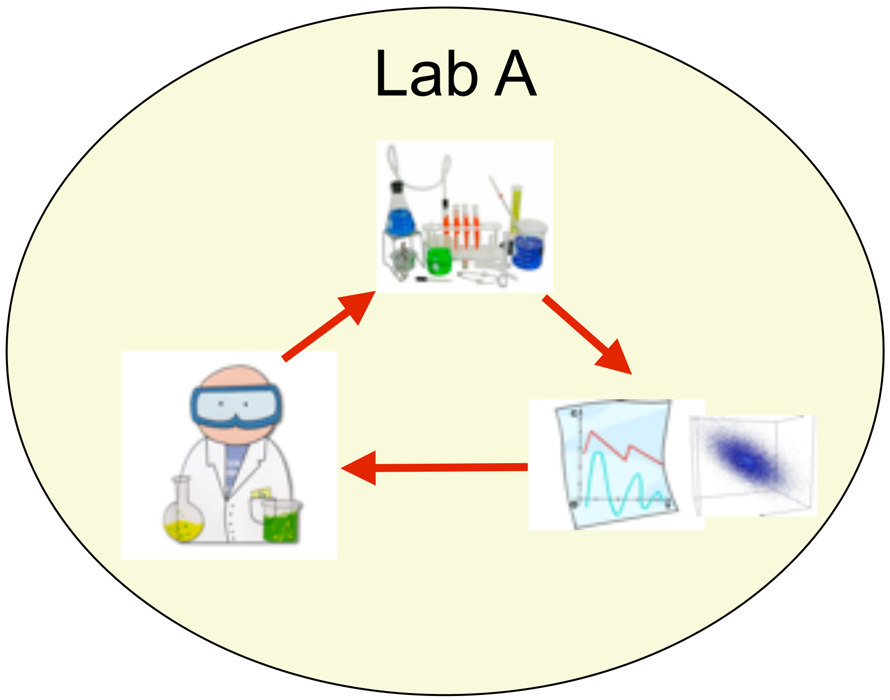
observe / infer → Discovery !

selective sensing

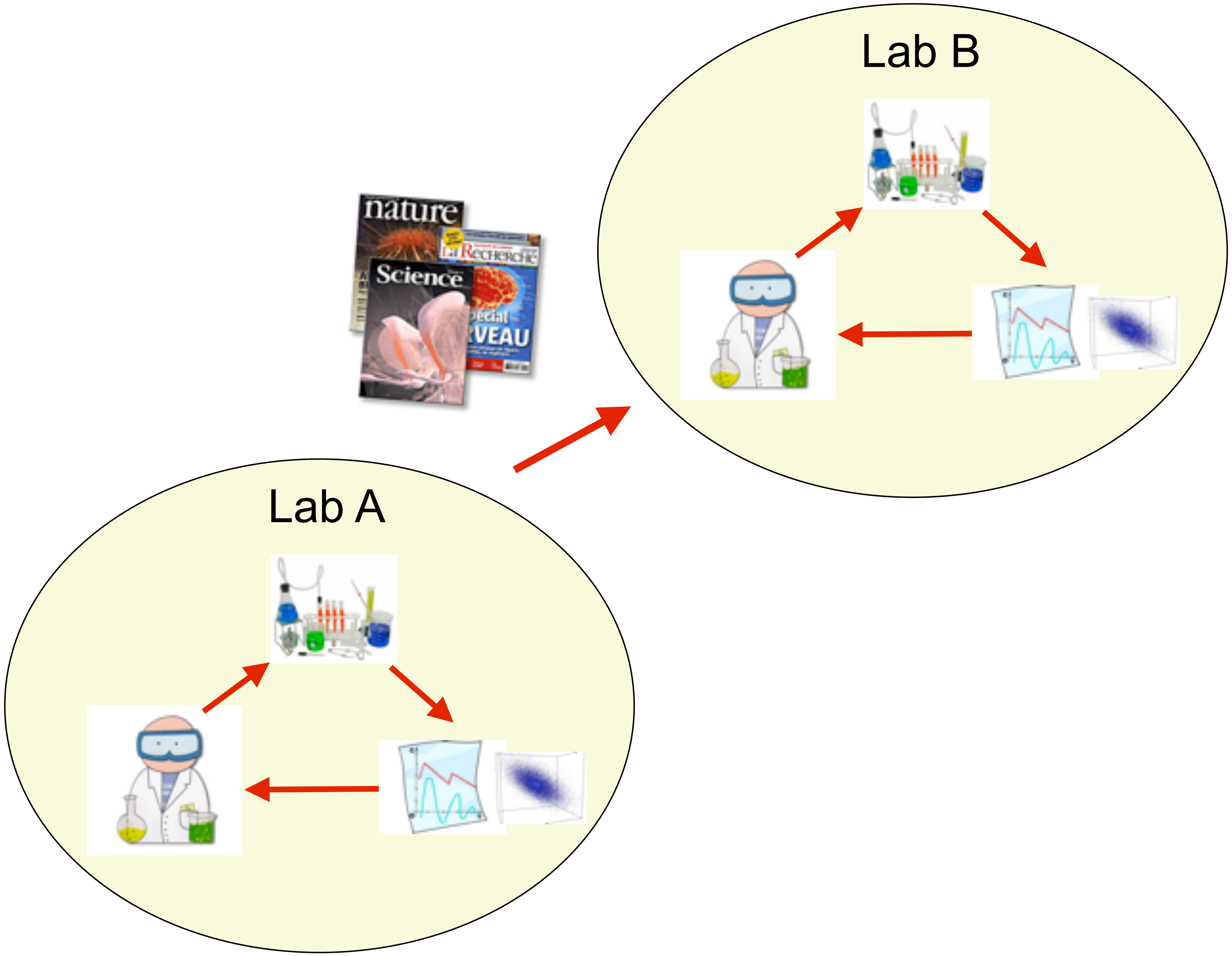
The Scientific Process in a Laboratory



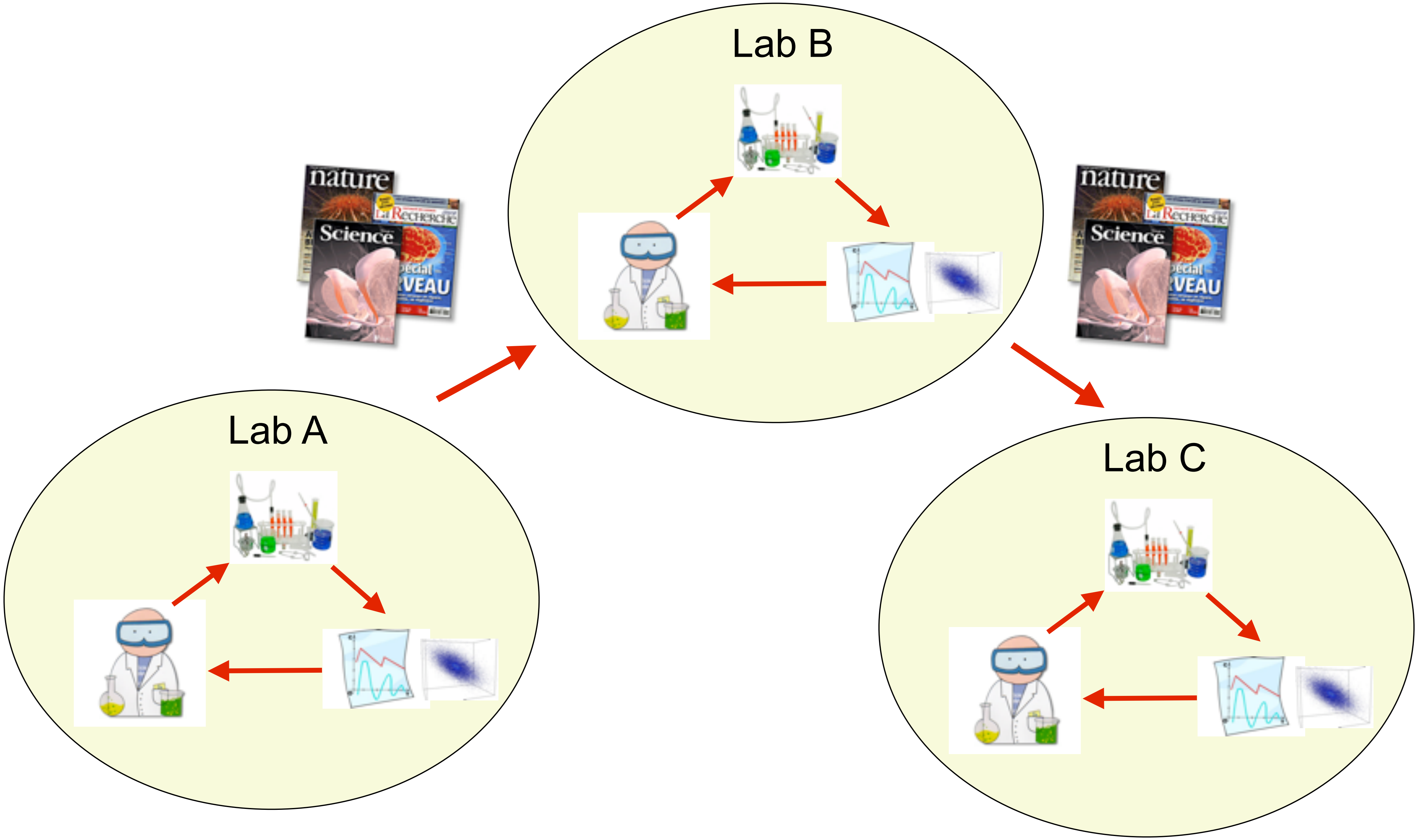
The Scientific Process at Large



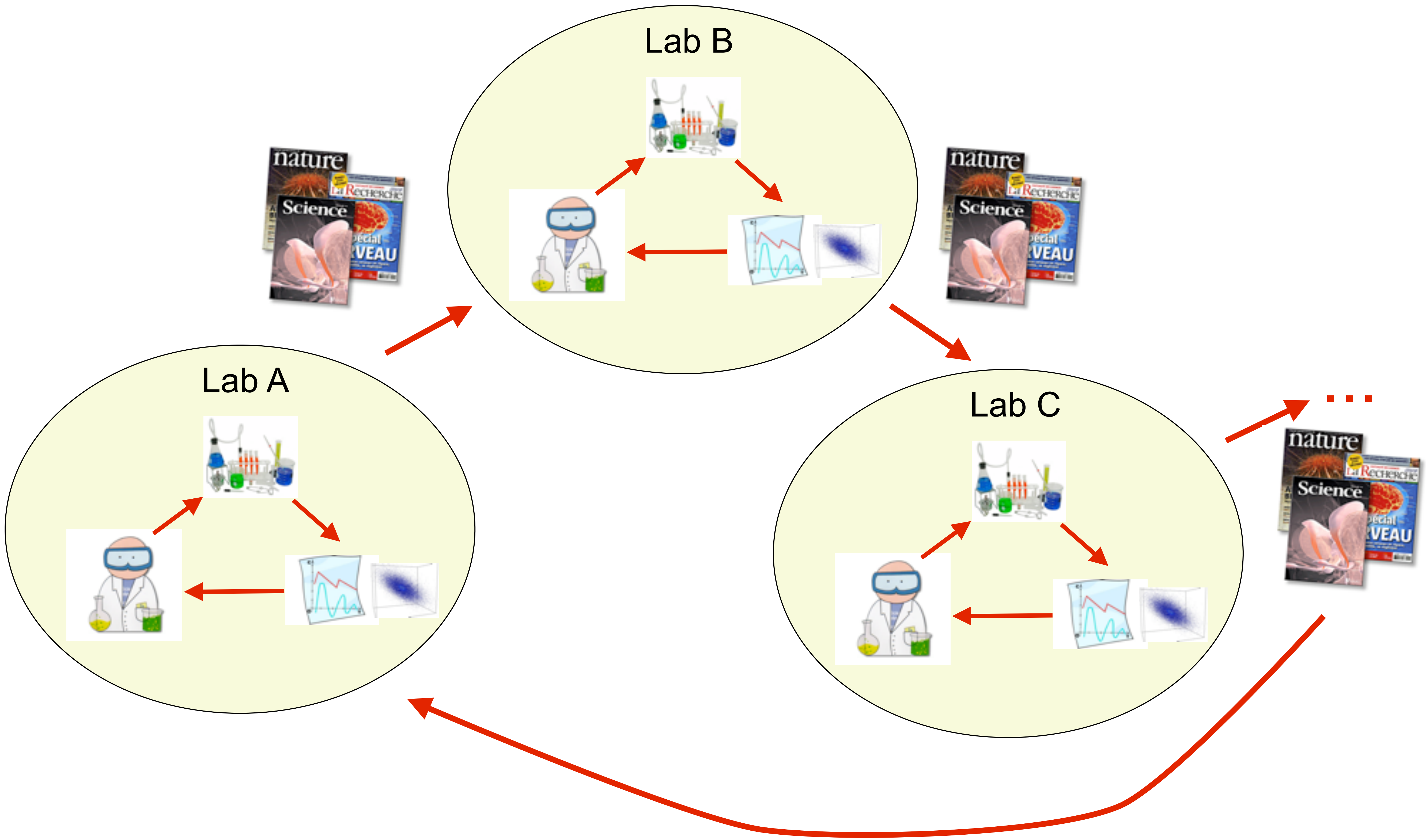
The Scientific Process at Large



The Scientific Process at Large



The Scientific Process at Large



Motivation: Inferring Biological Pathways

Paul Alhquist
(Molecular Virology)

Audrey Gasch
(Genetics)

Motivation: Inferring Biological Pathways

Paul Alhquist
(Molecular Virology)

Audrey Gasch
(Genetics)

virus

fruit fly

Motivation: Inferring Biological Pathways

virus

13,071 single-gene
knock-down cell strains

fruit fly

Paul Alhquist
(Molecular Virology)

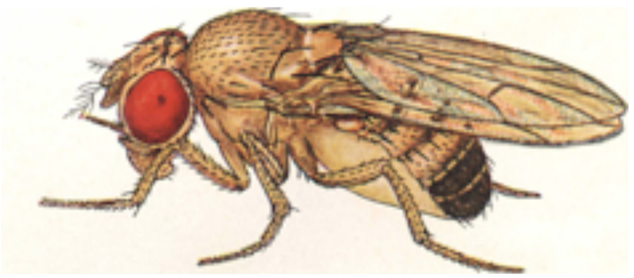
Audrey Gasch
(Genetics)

Motivation: Inferring Biological Pathways

Paul Alhquist
(Molecular Virology)

Audrey Gasch
(Genetics)

virus



fruit fly

13,071 single-gene
knock-down cell strains

infect each strain
with fluorescing virus

microwell
array

Motivation: Inferring Biological Pathways

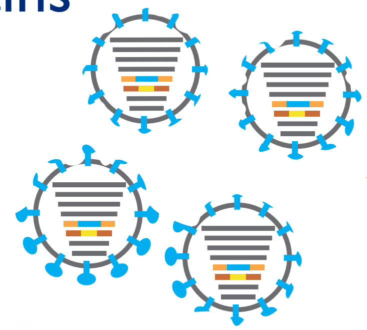
Paul Alhquist
(Molecular Virology)

Audrey Gasch
(Genetics)

virus

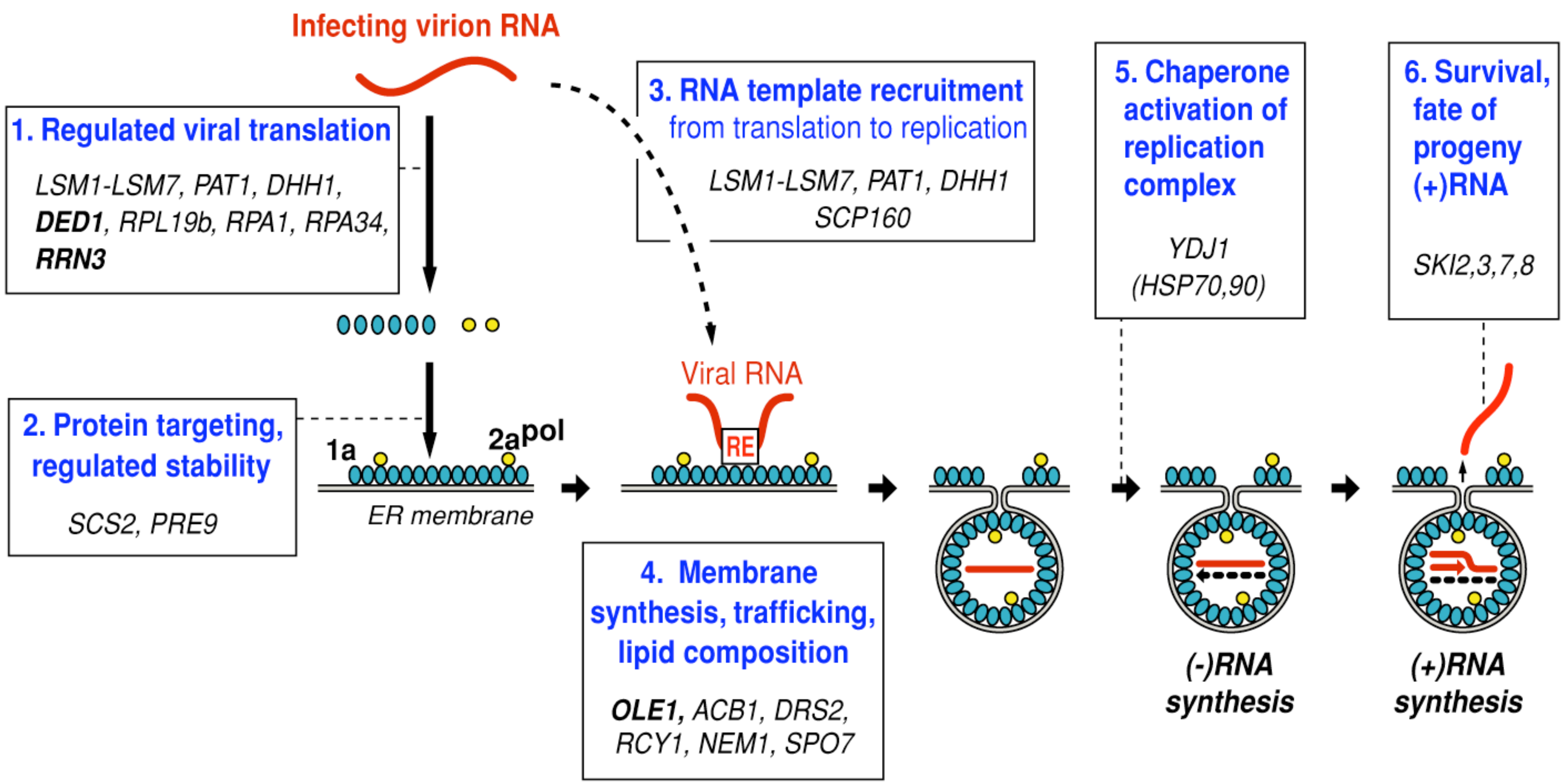
fruit fly

13,071 single-gene
knock-down cell strains



infect each strain
with fluorescing virus

microwell
array



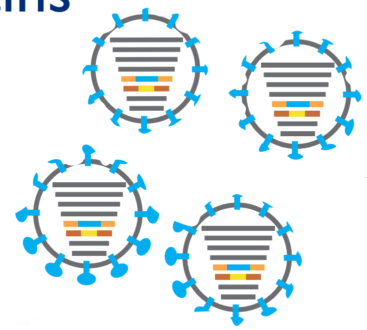
Motivation: Inferring Biological Pathways

Paul Alhquist
(Molecular Virology)

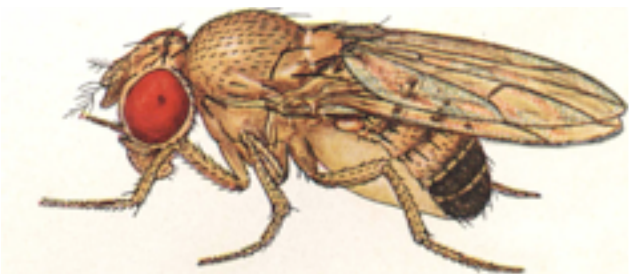
Audrey Gasch
(Genetics)

virus

13,071 single-gene
knock-down cell strains



infect each strain
with fluorescing virus



fruit fly

microwell
array

"Drosophila RNAi screen identifies host genes important for influenza virus replication," Nature 2008. How do they confidently determine the ~100 out of 13K genes hijacked for virus replication from extremely noisy data?

Motivation: Inferring Biological Pathways

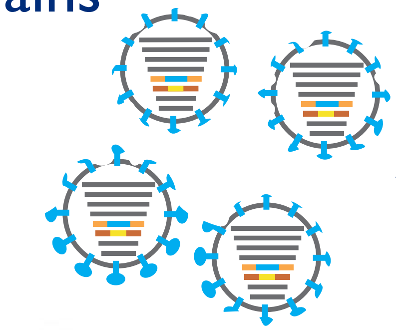
Paul Alhquist
(Molecular Virology)

Audrey Gasch
(Genetics)

virus

fruit fly

13,071 single-gene
knock-down cell strains



infect each strain
with fluorescing virus

microwell
array

"Drosophila RNAi screen identifies host genes important for influenza virus replication," Nature 2008. How do they confidently determine the ~100 out of 13K genes hijacked for virus replication from extremely noisy data?

Sequential Experimental Design:

Stage 1: assay all 13K strains, twice; keep all with significant fluorescence in one or both assays for 2nd stage (13K $\rightarrow$ 1K)

Stage 2: assay remaining 1K strains, 6-12 times; retain only those with statistically significant fluorescence (1K $\rightarrow$ 100)

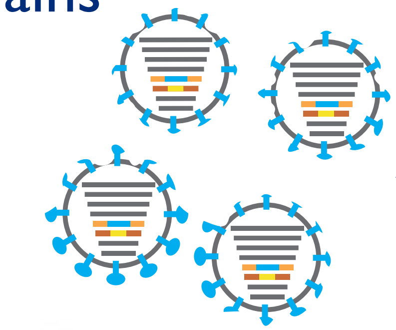
Motivation: Inferring Biological Pathways

Paul Alhquist
(Molecular Virology)

Audrey Gasch
(Genetics)

virus

13,071 single-gene
knock-down cell strains



infect each strain
with fluorescing virus

fruit fly

microwell
array

"Drosophila RNAi screen identifies host genes important for influenza virus replication," Nature 2008. How do they confidently determine the ~100 out of 13K genes hijacked for virus replication from extremely noisy data?

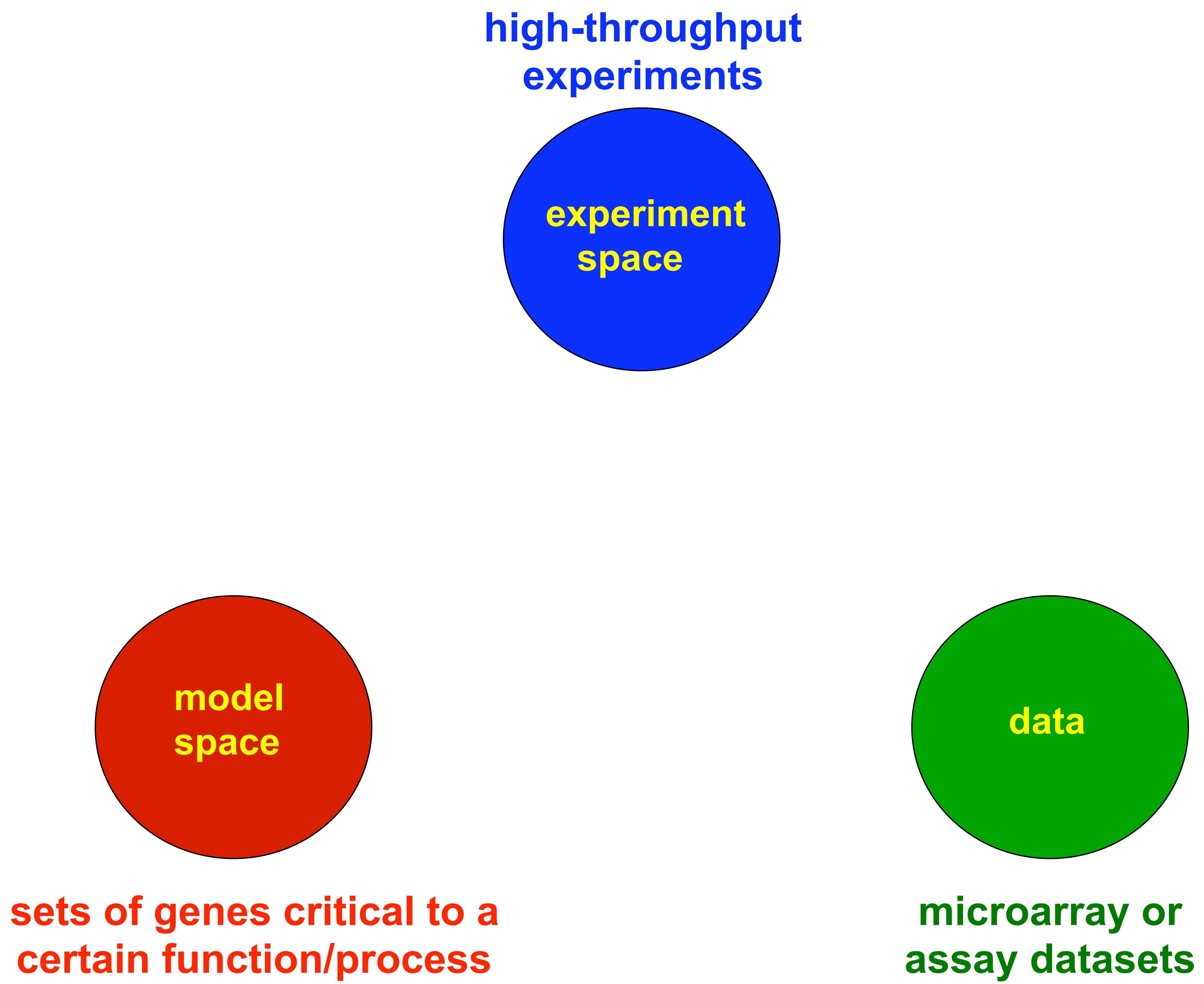
Sequential Experimental Design:

Stage 1: assay all 13K strains, twice; keep all with significant fluorescence in one or both assays for 2nd stage (13K $\rightarrow$ 1K)

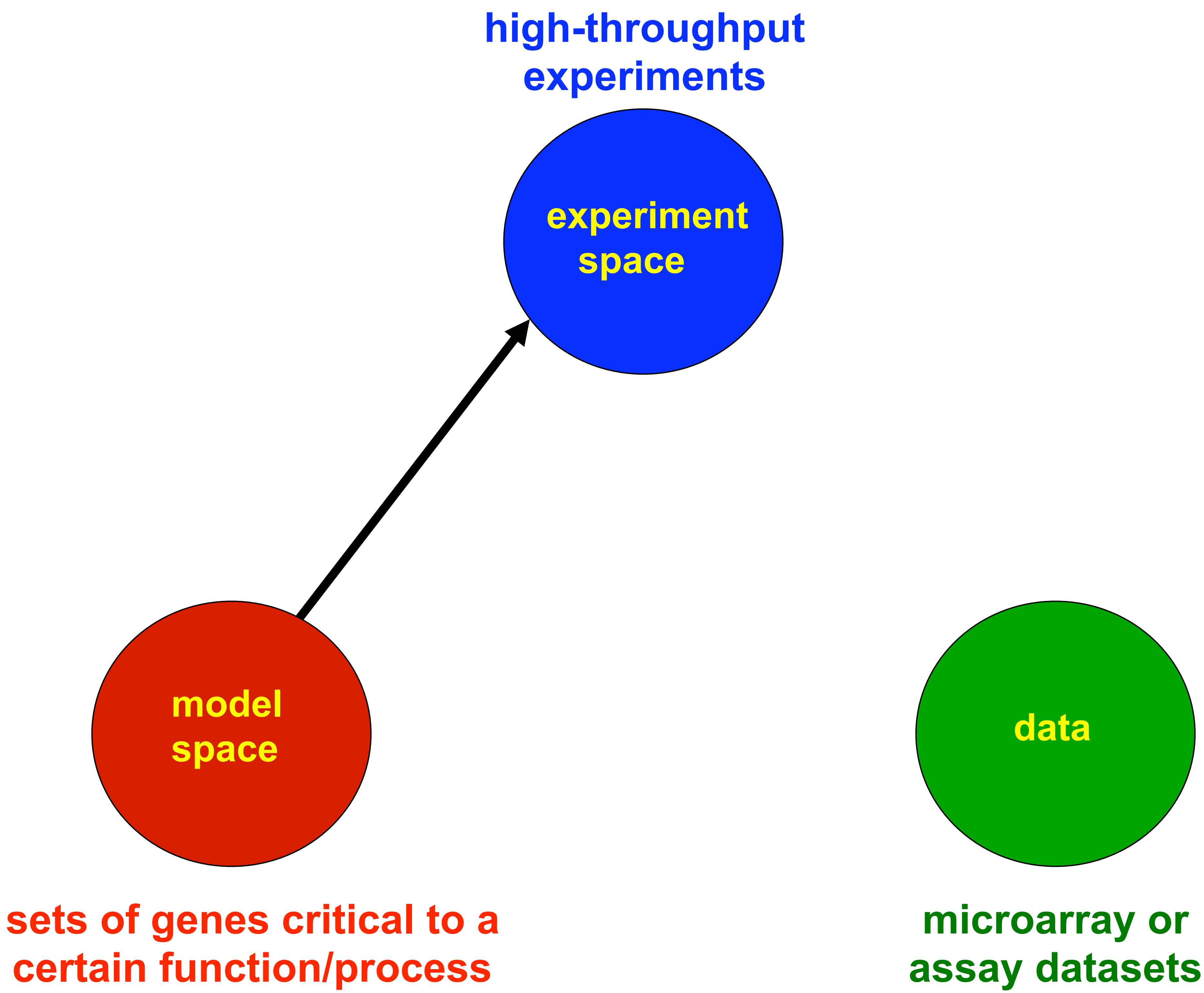
Stage 2: assay remaining 1K strains, 6-12 times; retain only those with statistically significant fluorescence (1K $\rightarrow$ 100)

vastly more efficient that replicating all 13K experiments many times

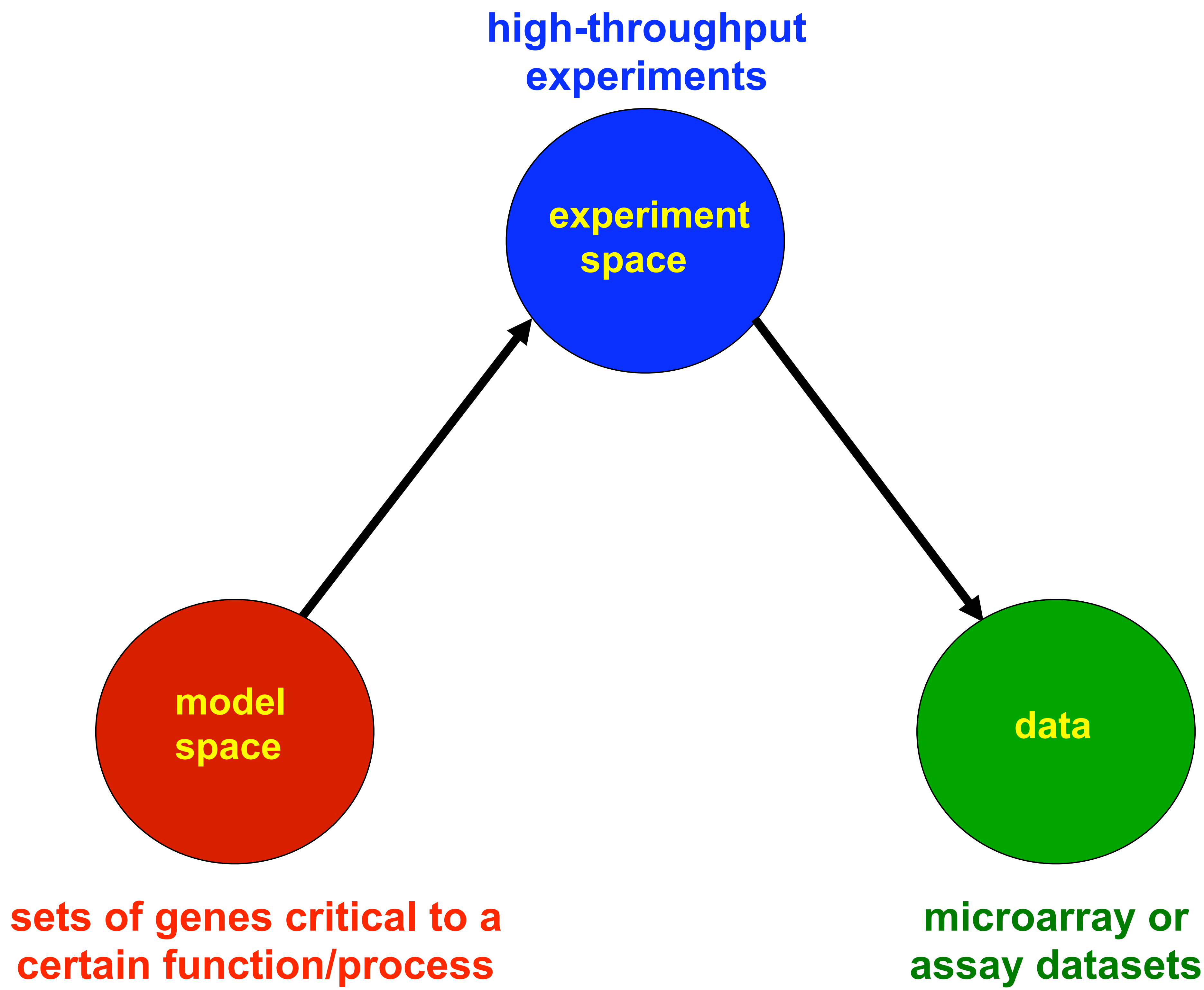
Feedback from Data Analysis to Data Collection



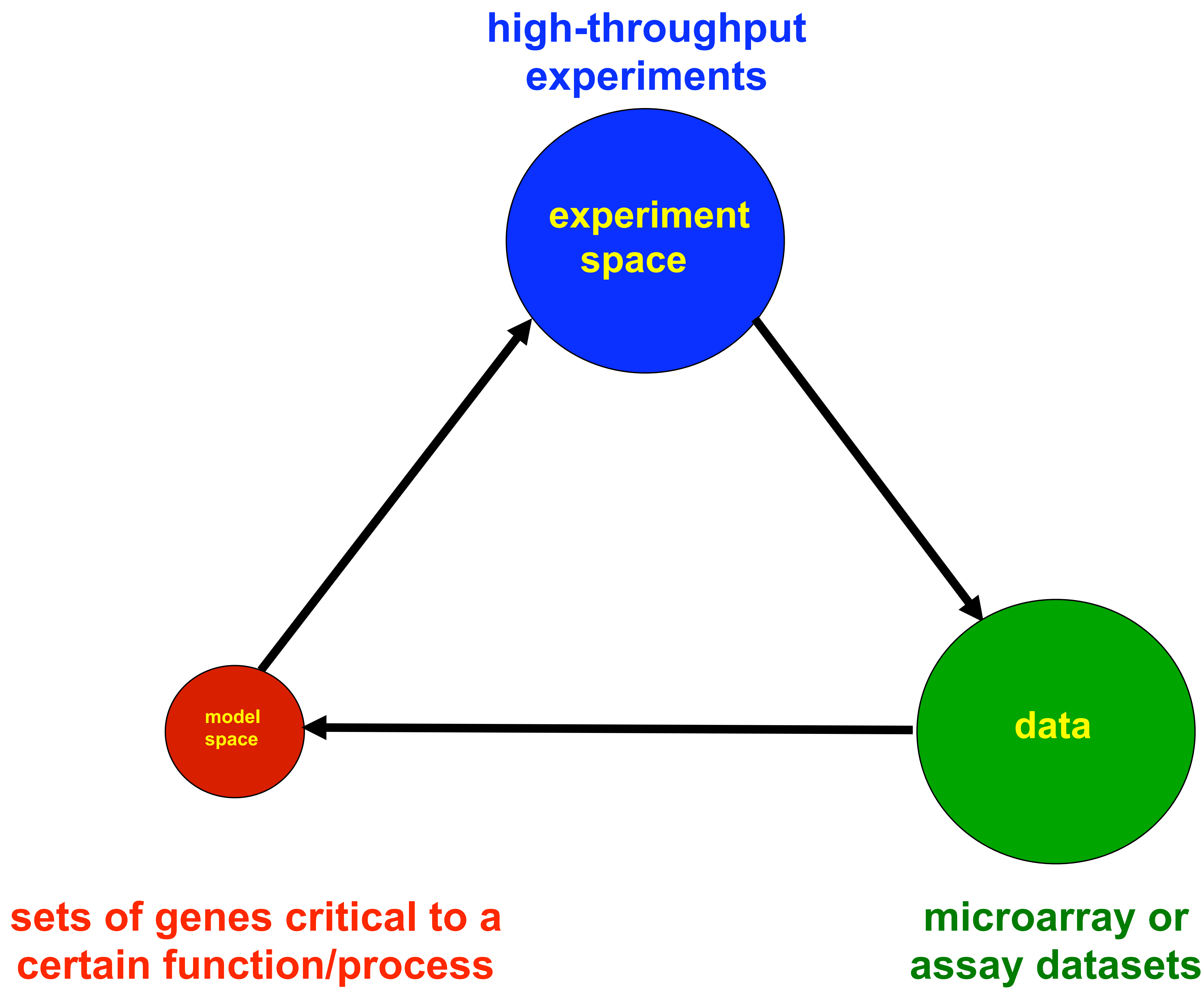
Feedback from Data Analysis to Data Collection



Feedback from Data Analysis to Data Collection



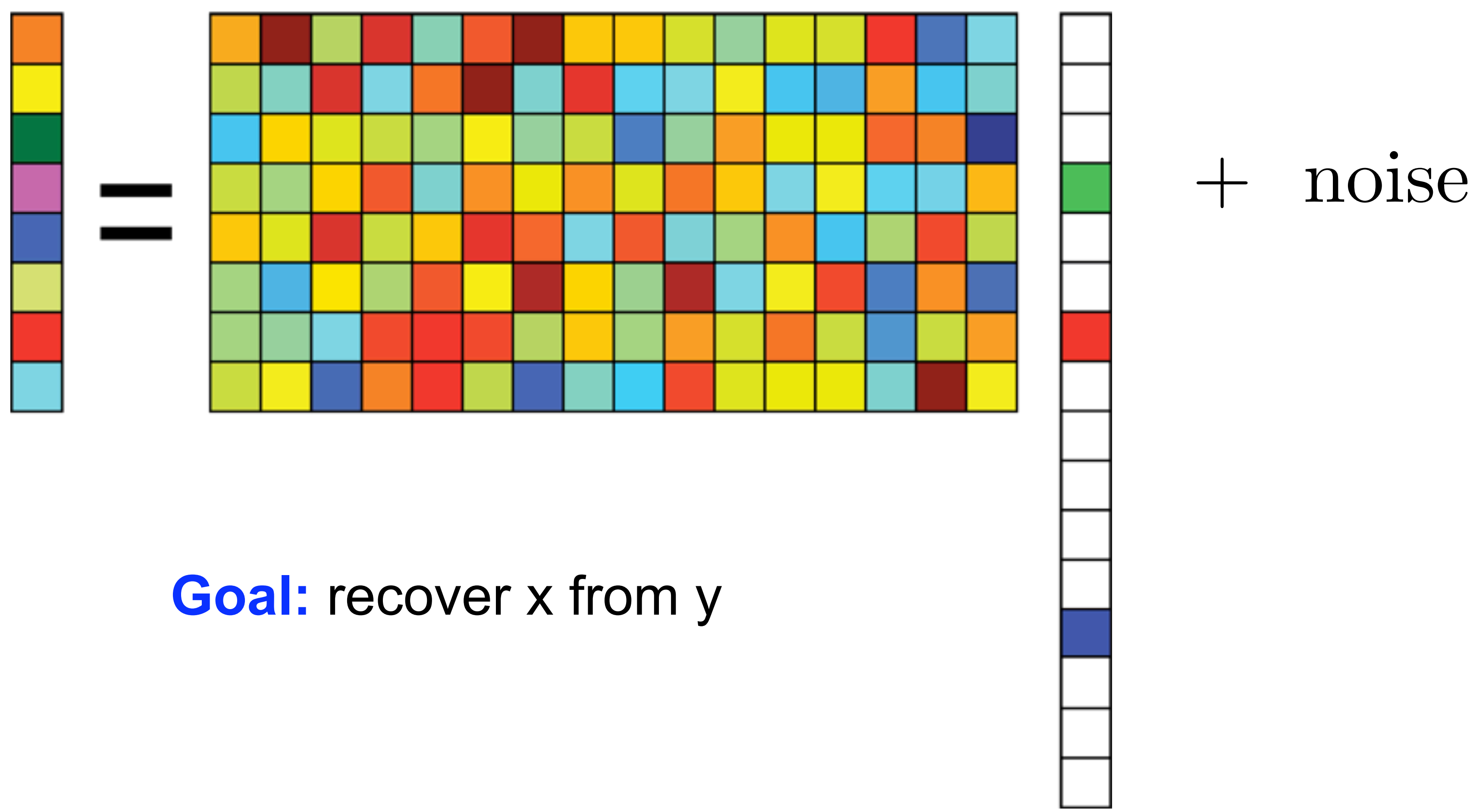
Feedback from Data Analysis to Data Collection



Adaptive Sensing for Sparse Recovery

(image reconstruction, compressed sensing, inverse problems)

$$y \;=\; A\,x \;+\; w \text{ , with } A \in \mathbb{R}^{m\times n},\ x \in \mathbb{R}^n \text{ (but sparse)},\ w \sim \mathcal{N}(0, I)$$

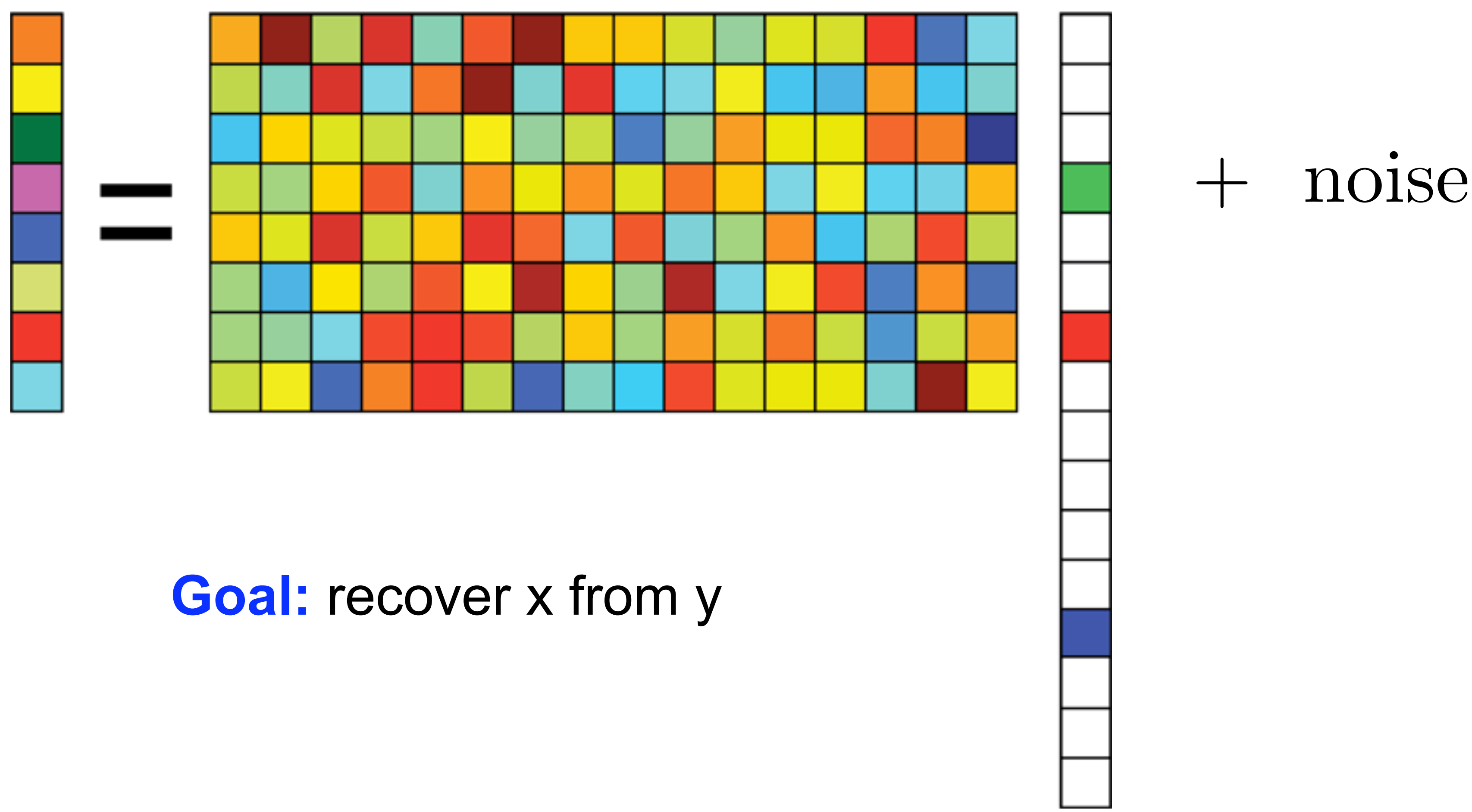


Goal: recover x from y

Adaptive Sensing for Sparse Recovery

(image reconstruction, compressed sensing, inverse problems)

$$y \;=\; A\,x \;+\; w \;, \text{ with } A \in \mathbb{R}^{m\times n},\ x \in \mathbb{R}^n \text{ (but sparse)},\ w \sim \mathcal{N}(0, I)$$



Goal: recover x from y

Is sequentially designing (rows of) A advantageous ?

Motivation: Randomized Experiments

$\widetilde{y}$ $=$  $\times$  $+$ 

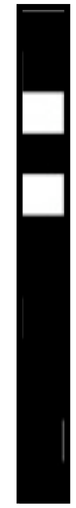
$k \times 1$

sparse signal

noise

indirect (randomized) measurement

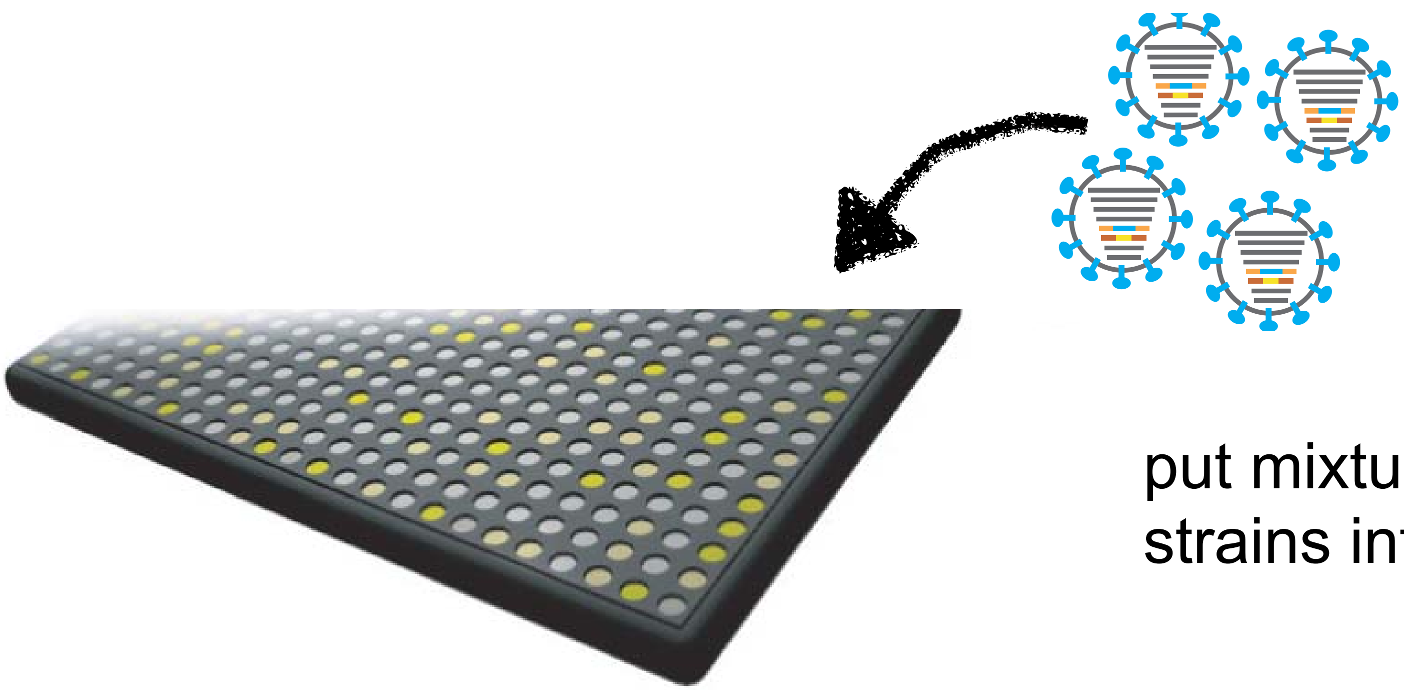
Motivation: Randomized Experiments

$\underset{k \times 1}{\widetilde{y}} =$  $\times$  $+$ 

sparse signal

noise

indirect (randomized) measurement



put mixtures of single-deletion strains into each well

Signal Processing Gear

back to shop

y = phi x (Dark T-Shirt)

$18.99

Fit: Standard

Not too tight, not too loose.

Fabric Thickness:

Thin Thick

1. Color: (Charcoal)
2. Size: Large Size Chart
3. Qty: 1

ADD TO CART

AVAILABILITY: In Stock.
Product Number: 030-469487567

Like Sign Up to see what your friends like.

Share |

A visual representation of the math behind compressive sensing

Look cool without breaking the bank. Our durable, high-quality, pre-shrunk 100% cotton t-shirt is what to wear when you want to go comfortably casual. Preshrunk, durable and guaranteed.

- 5.6 oz. 100% cotton
- Standard fit

Other items by Signal Processing Gear:

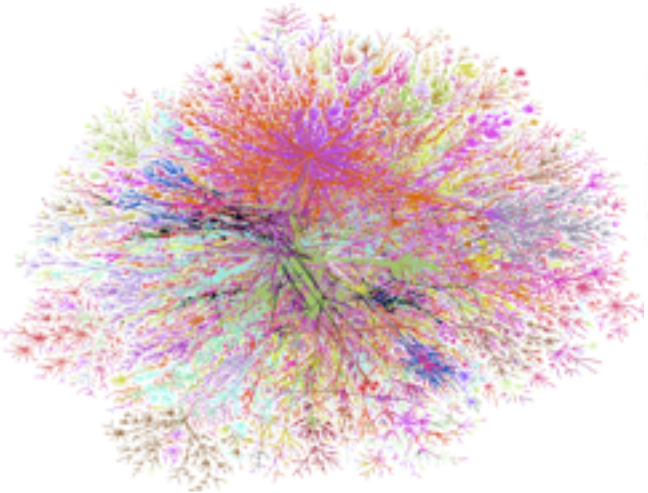
y = phi x (Mug)

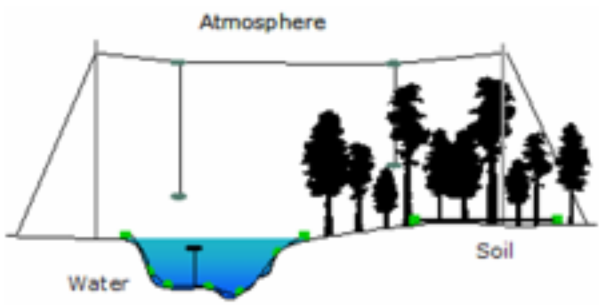
y = phi x (Large Mug)

y = phi x (Light T-Shirt)

Sensing and Inference in Large Networked Systems

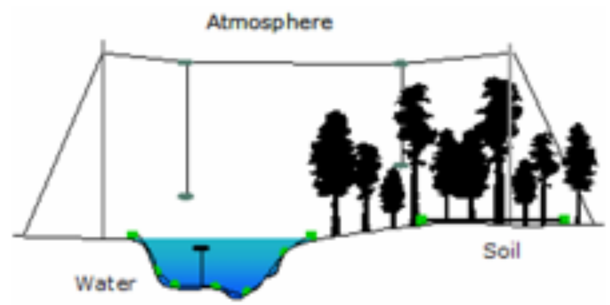
Sensing and Inference in Large Networked Systems





Technological Networks
(Internet Mapping Project, US power grid, UCLA CENS)

Sensing and Inference in Large Networked Systems



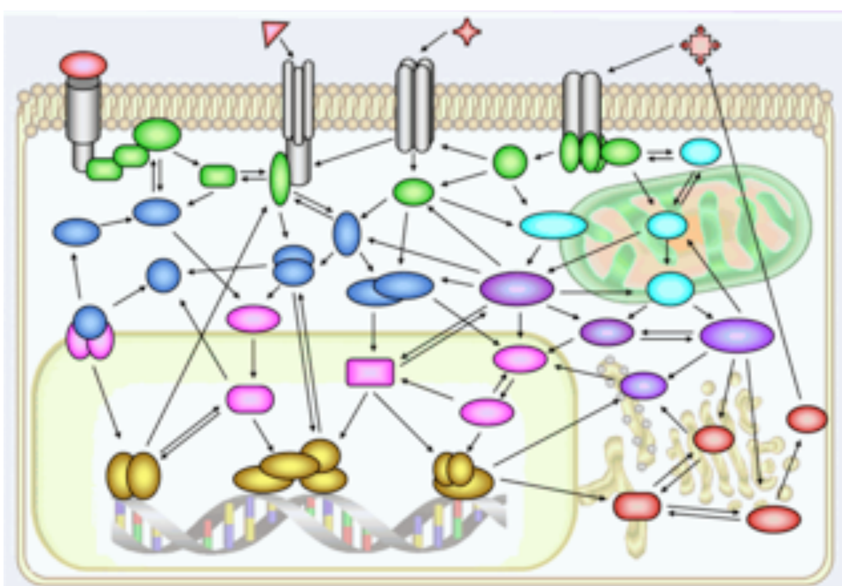
Technological Networks
(Internet Mapping Project, US power grid, UCLA CENS)

Social Networks

Sensing and Inference in Large Networked Systems

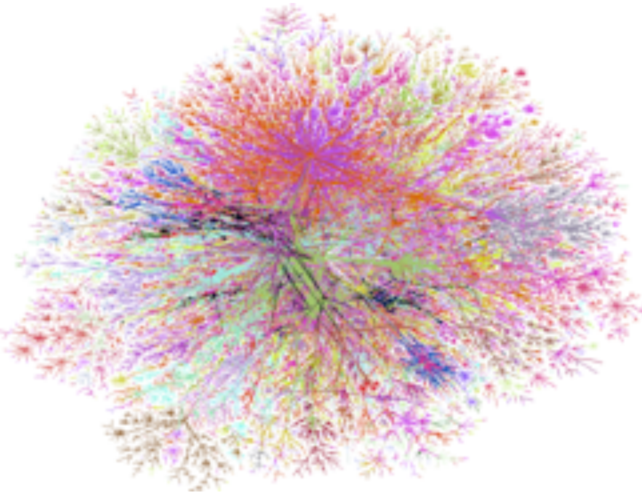
Technological Networks
(Internet Mapping Project, US power grid, UCLA CENS)

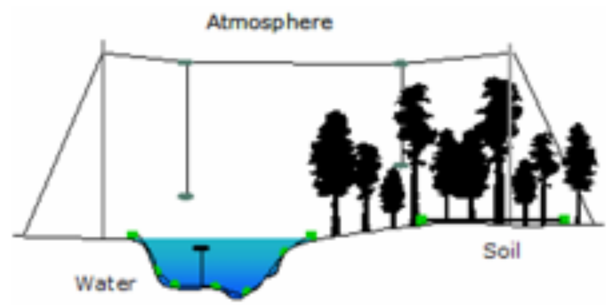
Social Networks



Biological Networks
(JMDBase)

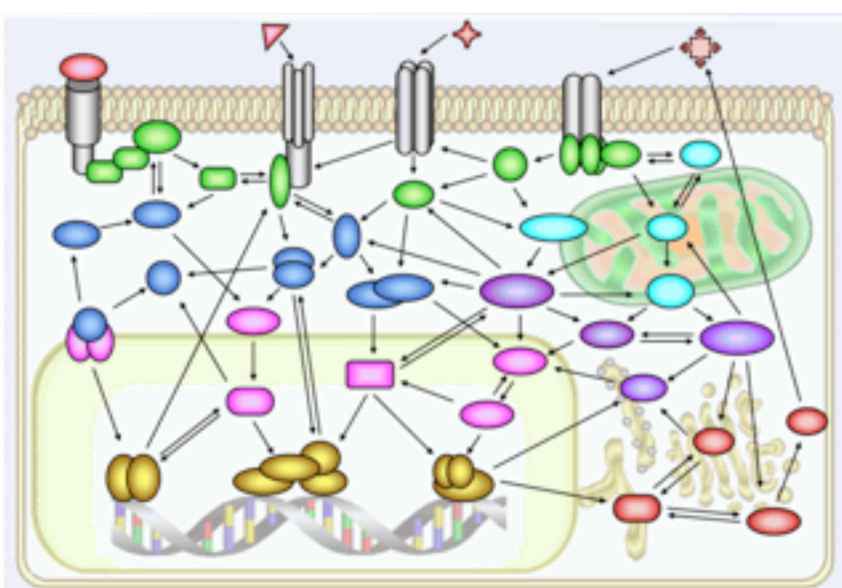
Sensing and Inference in Large Networked Systems



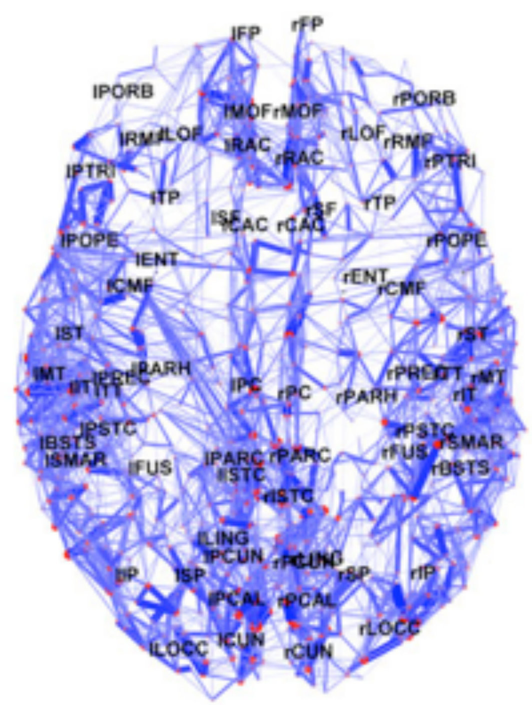


Technological Networks
(Internet Mapping Project, US power grid, UCLA CENS)

Social Networks



Biological Networks
(JMDBase)

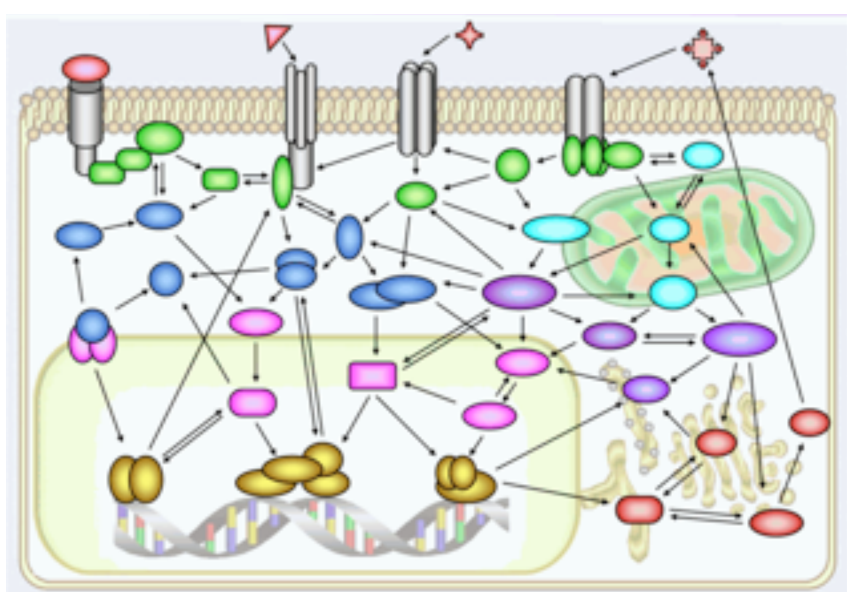


Brain Networks
(Worsley et al, 2005)

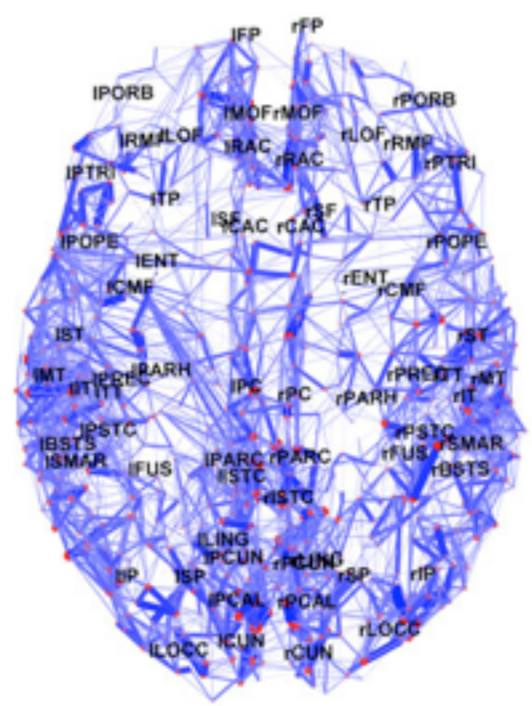
Sensing and Inference in Large Networked Systems

Technological Networks
(Internet Mapping Project, US power grid, UCLA CENS)

Social Networks



Biological Networks
(JMDBase)



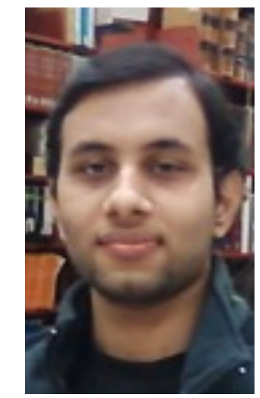
Brain Networks
(Worsley et al, 2005)

Challenges:

- Inferring structure & function of the system
- Optimized design & resource allocation
- Pattern analysis & anomaly detection

Network Structure and Clustering

Complex systems are not defined by the independent functions of individual components, rather they depend on the orchestrated interactions of these elements.



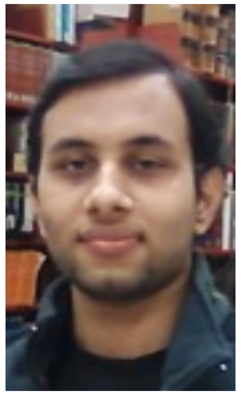
Gautam Dasarathy

Brian Eriksson

Network Structure and Clustering

Complex systems are not defined by the independent functions of individual components, rather they depend on the orchestrated interactions of these elements.

Network(s) of interactions can be revealed via **clustering** based on measured features



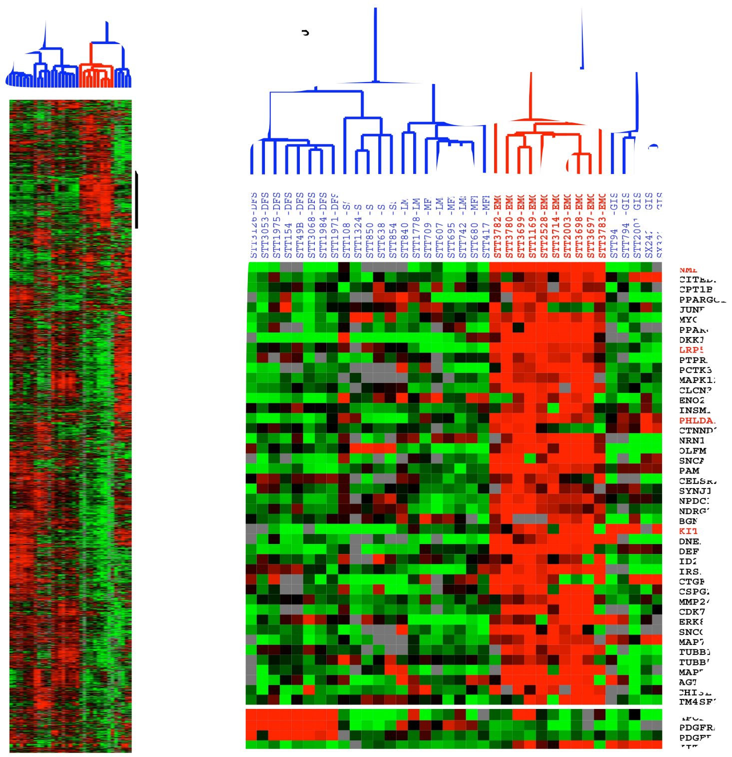
Gautam Dasarathy

Brian Eriksson

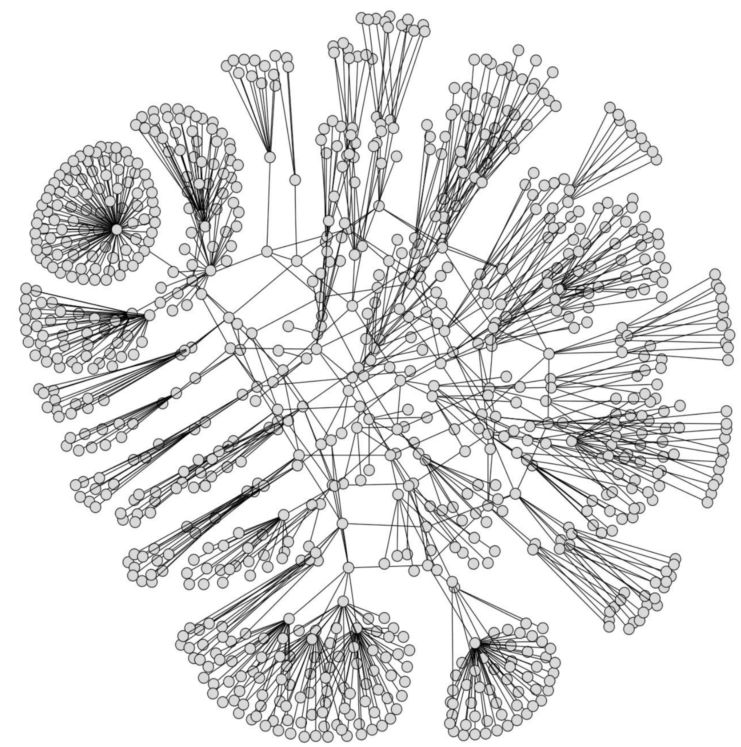
Network Structure and Clustering

Complex systems are not defined by the independent functions of individual components, rather they depend on the orchestrated interactions of these elements.

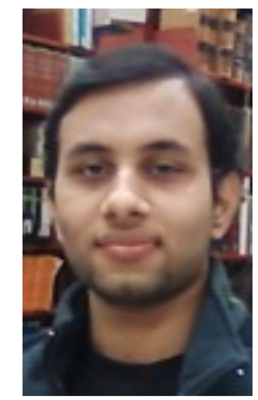
Network(s) of interactions can be revealed via **clustering** based on measured features



genes and expression/ interaction profiles



network routers and traffic/distance profiles



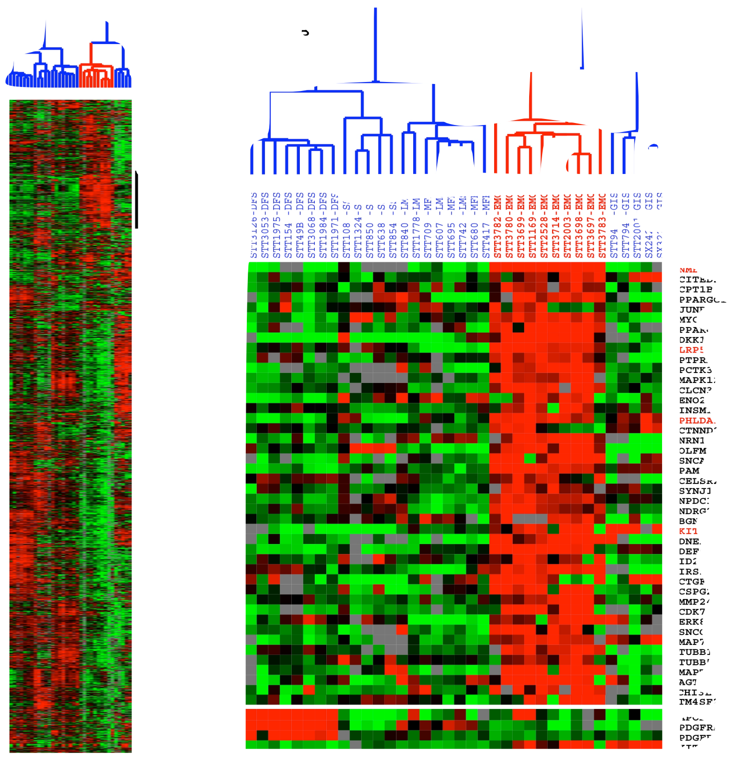
Gautam Dasarathy

Brian Eriksson

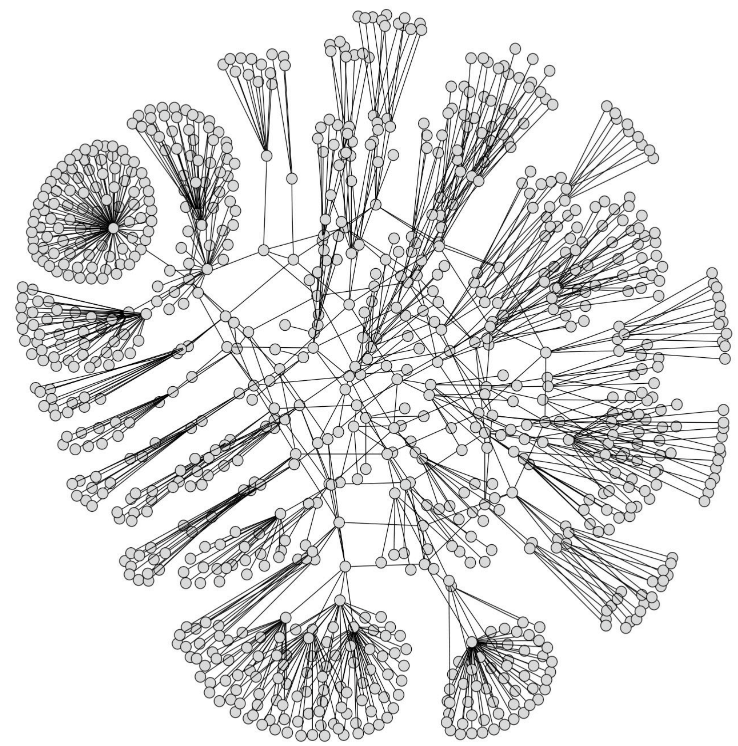
Network Structure and Clustering

Complex systems are not defined by the independent functions of individual components, rather they depend on the orchestrated interactions of these elements.

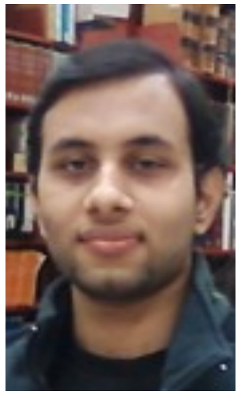
Network(s) of interactions can be revealed via **clustering** based on measured features



genes and expression/
interaction profiles



network routers and
traffic/distance profiles



Gautam
Dasarathy

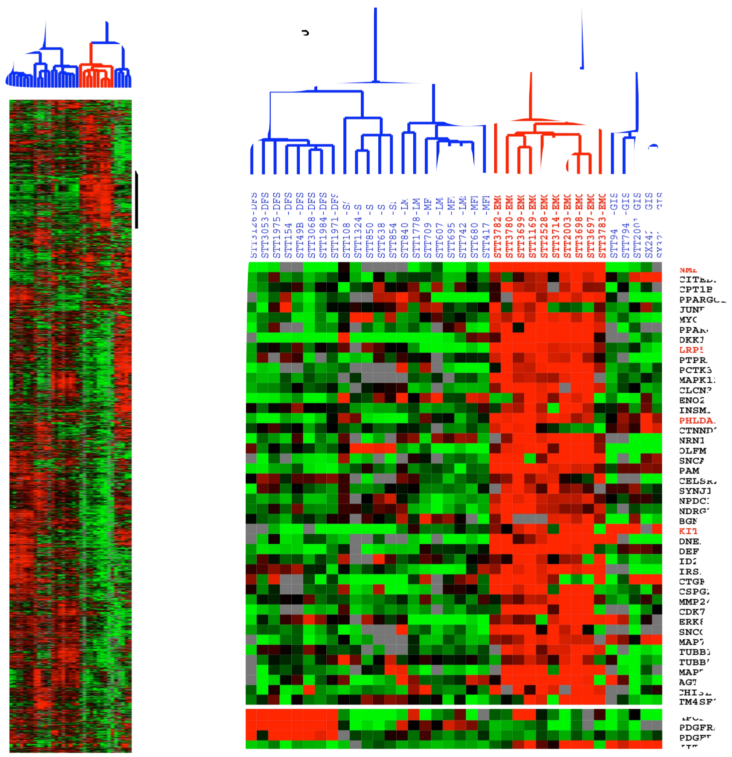
Brian
Eriksson

Similarity-Based Clustering: Each component (gene/router) has an associated feature (measurement profile). Components can be clustered based on feature similarities.

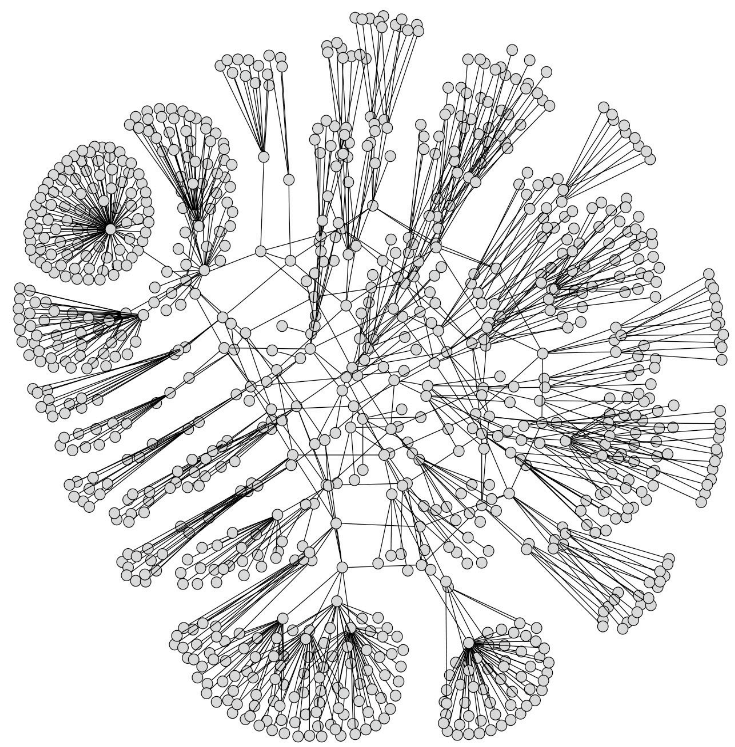
Network Structure and Clustering

Complex systems are not defined by the independent functions of individual components, rather they depend on the orchestrated interactions of these elements.

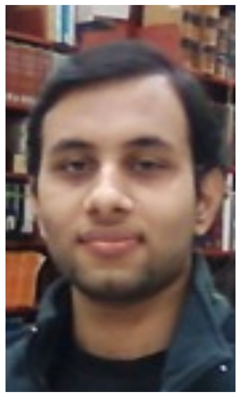
Network(s) of interactions can be revealed via **clustering** based on measured features



genes and expression/
interaction profiles



network routers and
traffic/distance profiles



Gautam
Dasarathy

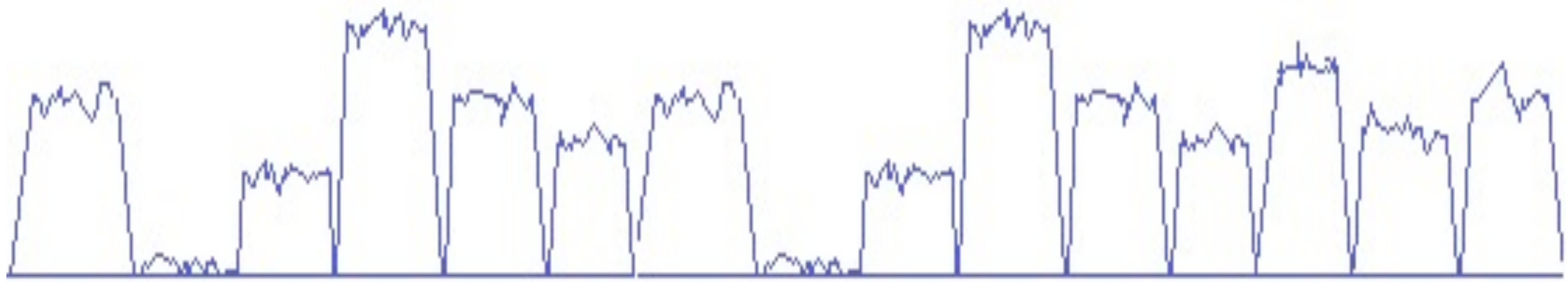
Brian
Eriksson

Similarity-Based Clustering: Each component (gene/router) has an associated feature (measurement profile). Components can be clustered based on feature similarities.

Recent Result: A sequential method for selecting "informative" similarities that produces accurate clusters from as few as $3N \log N$ similarities.

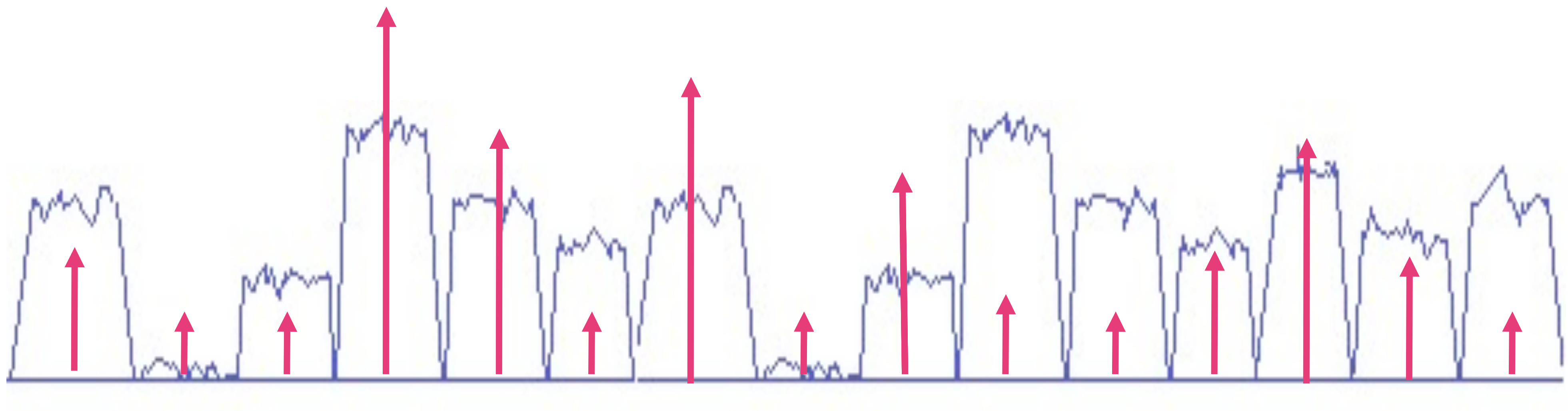
Cognitive Radio Spectrum Sensing

"primary" users have preference over "secondary" users



Cognitive Radio Spectrum Sensing

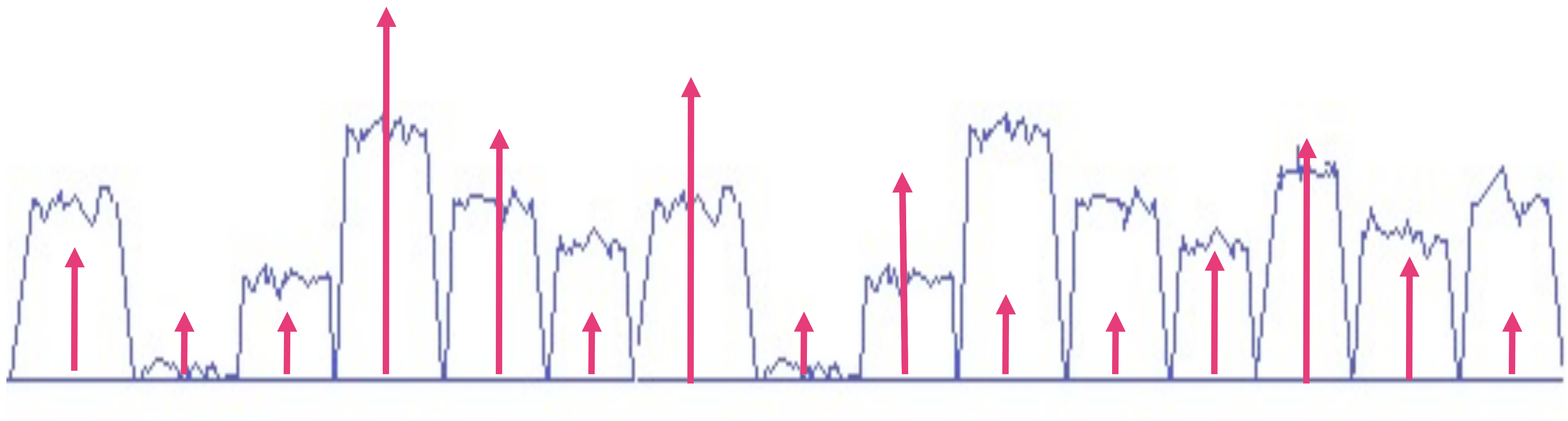
"primary" users have preference over "secondary" users



most channels occupied by primary users, but they come and go in unpredictable manner. Secondary users "sense" spectrum to find an unoccupied channel

Cognitive Radio Spectrum Sensing

"primary" users have preference over "secondary" users



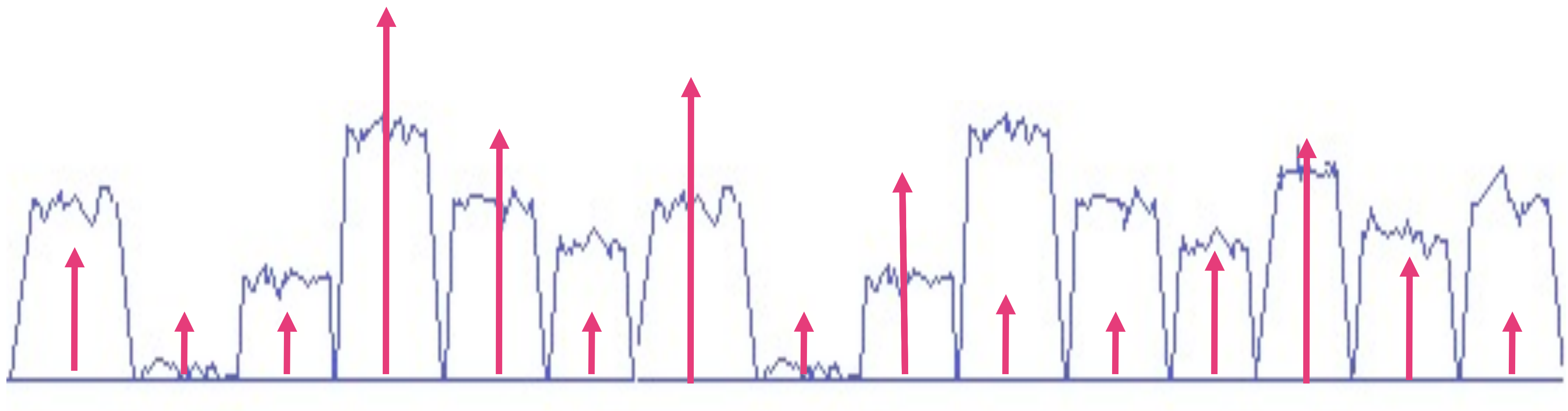
most channels occupied by primary users, but they come and go in unpredictable manner. Secondary users "sense" spectrum to find an unoccupied channel

Goal: Find open channel(s) as quickly as possible. Two approaches:

1) listen to each channel for a fixed amount of time and make decision
2) listen to each channel for a data-adaptive amount of time to make decisions as quickly as possible

Cognitive Radio Spectrum Sensing

"primary" users have preference over "secondary" users



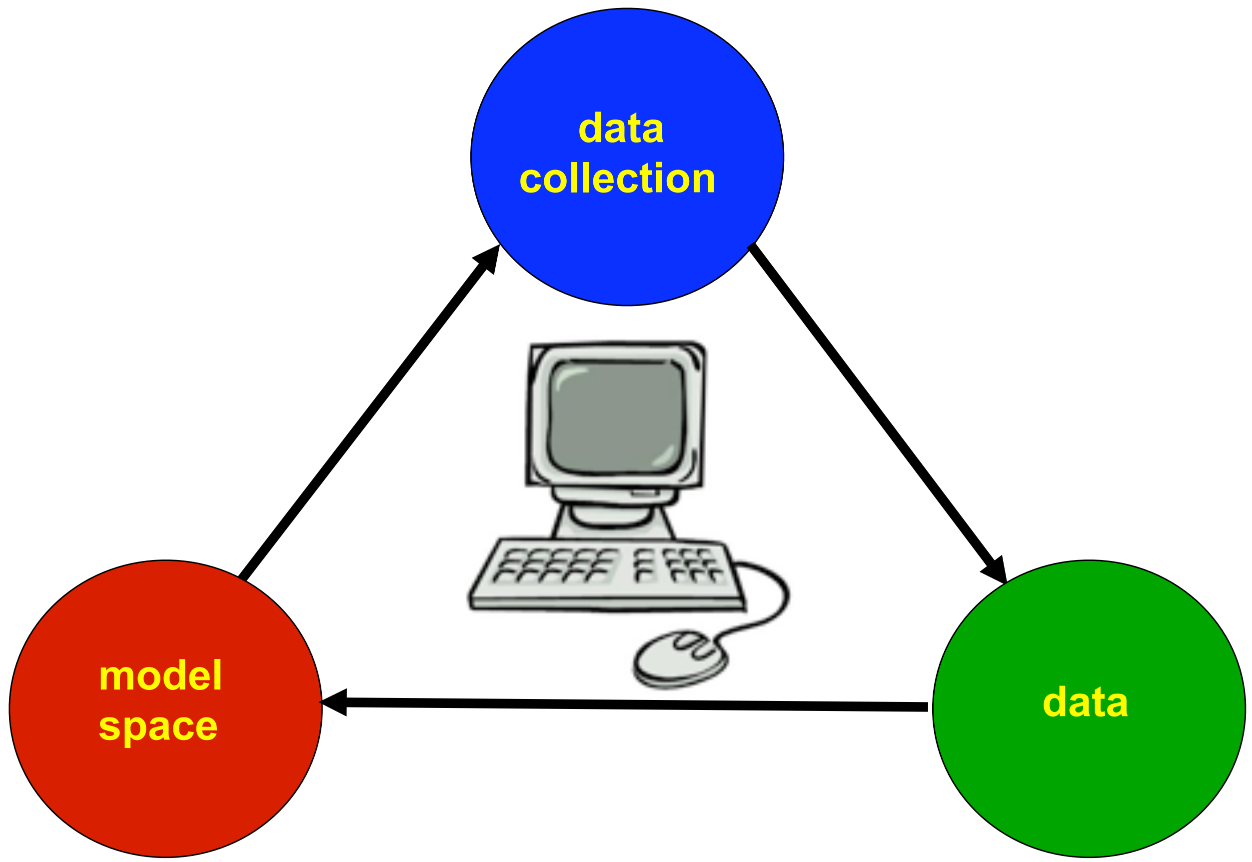
most channels occupied by primary users, but they come and go in unpredictable manner. Secondary users "sense" spectrum to find an unoccupied channel

Goal: Find open channel(s) as quickly as possible. Two approaches:

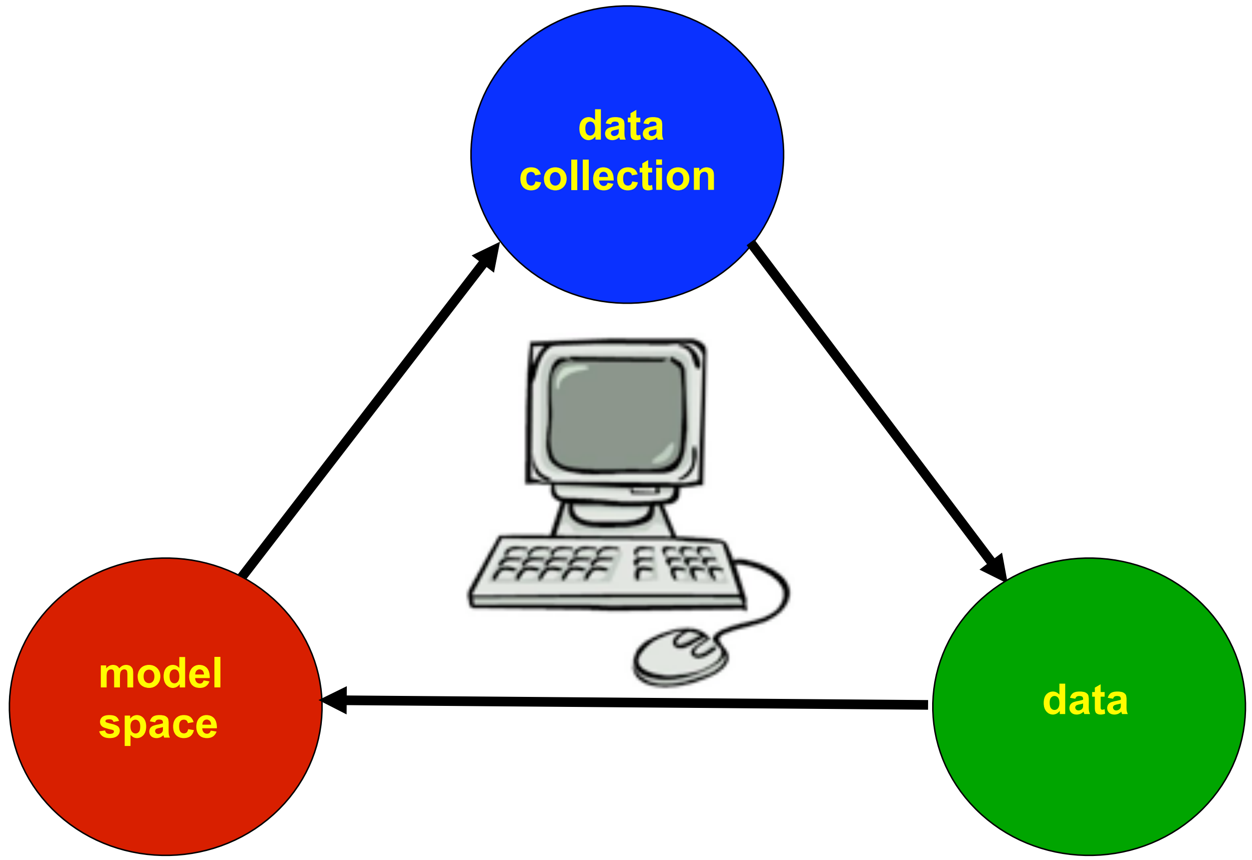
1) listen to each channel for a fixed amount of time and make decision
2) listen to each channel for a data-adaptive amount of time to make decisions as quickly as possible

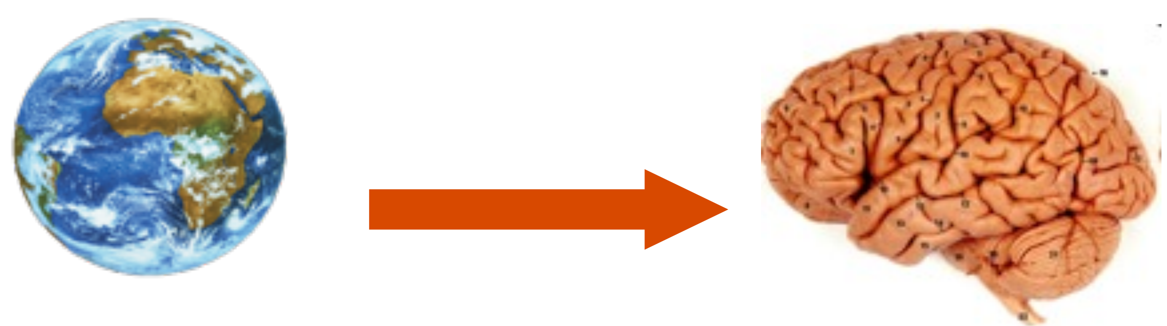
adaptive spectrum sensing is significantly more time-efficient than fixed sensing

Active Learning in Machines and Humans



Active Learning in Machines and Humans

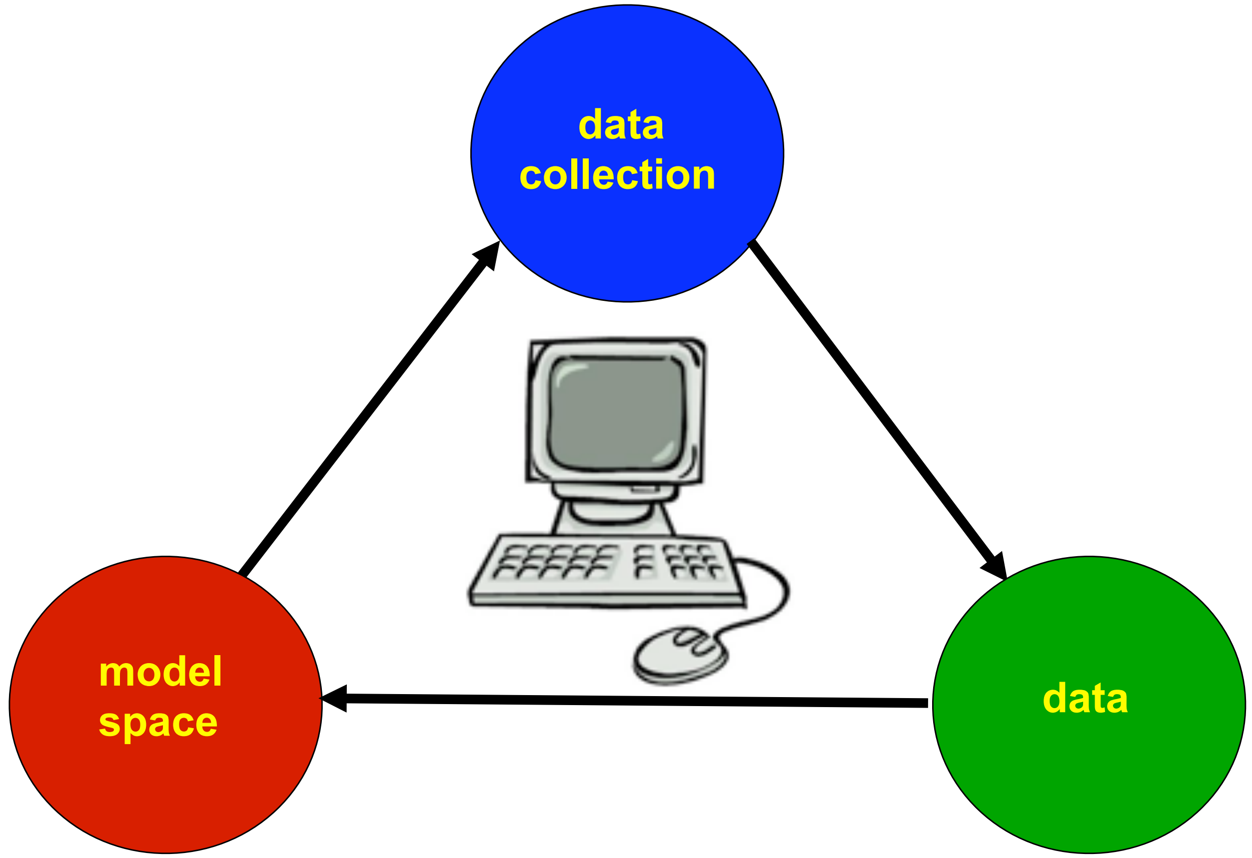


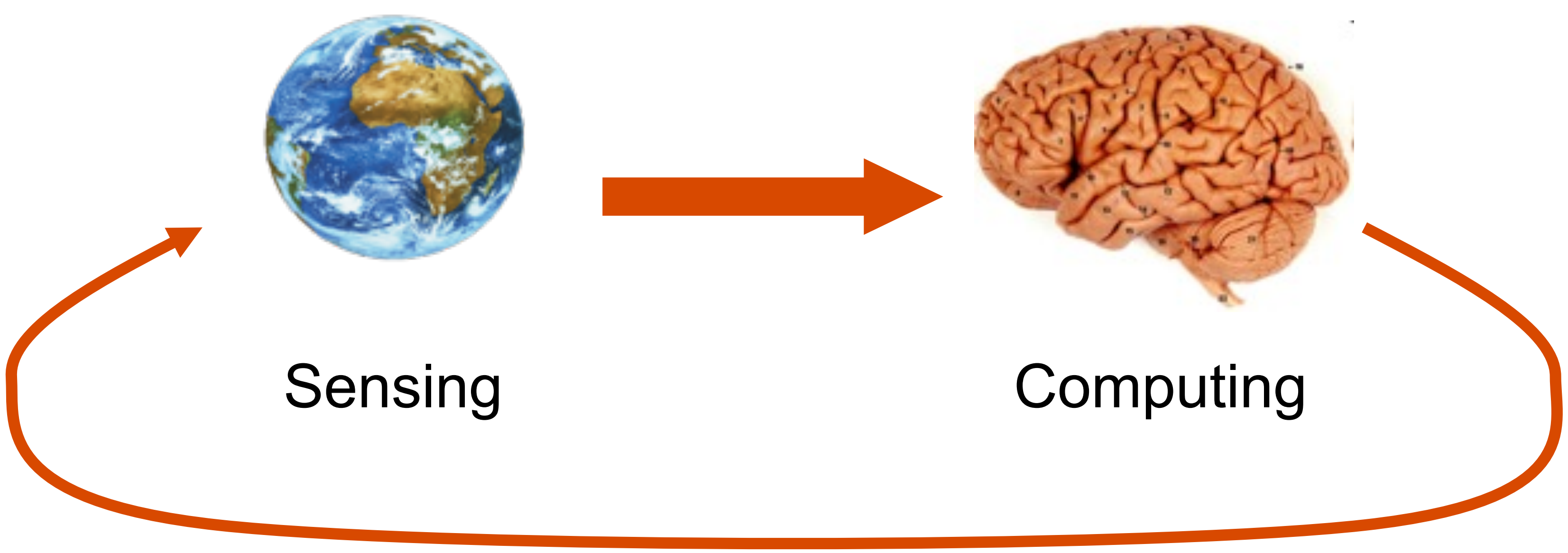


Sensing

Computing

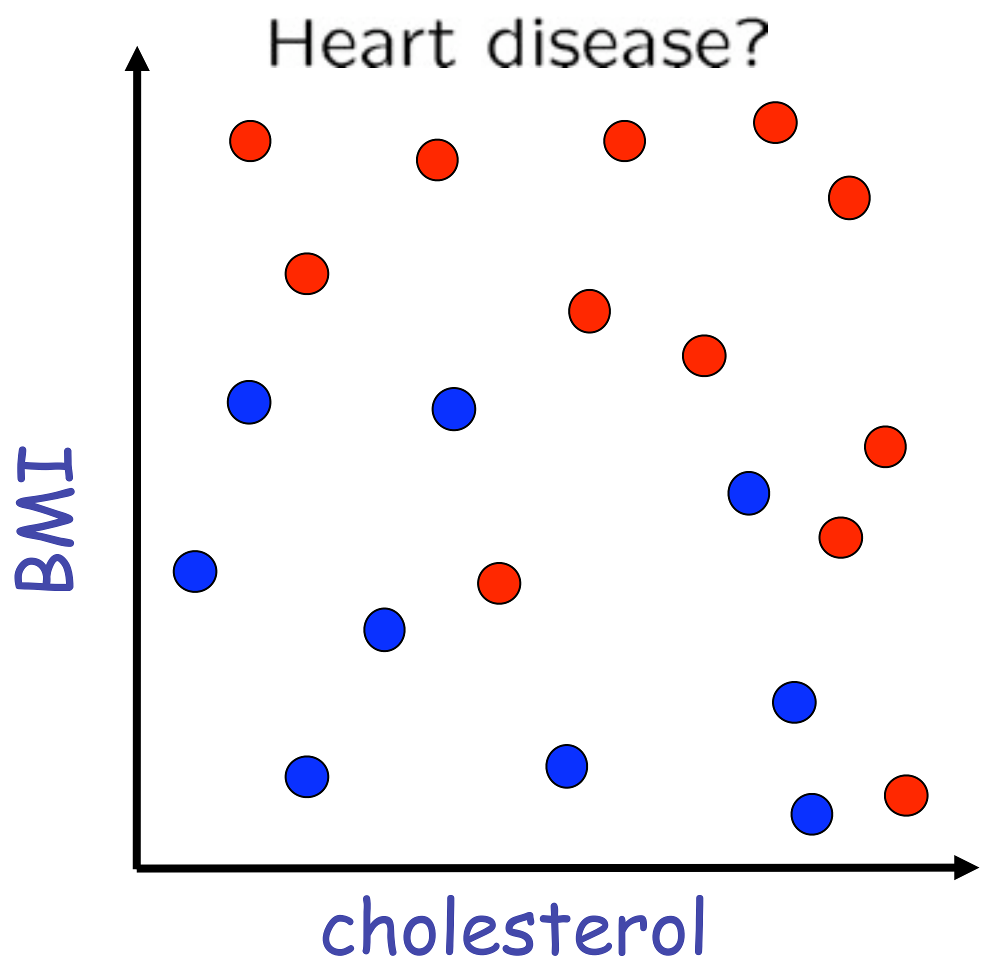
Active Learning in Machines and Humans





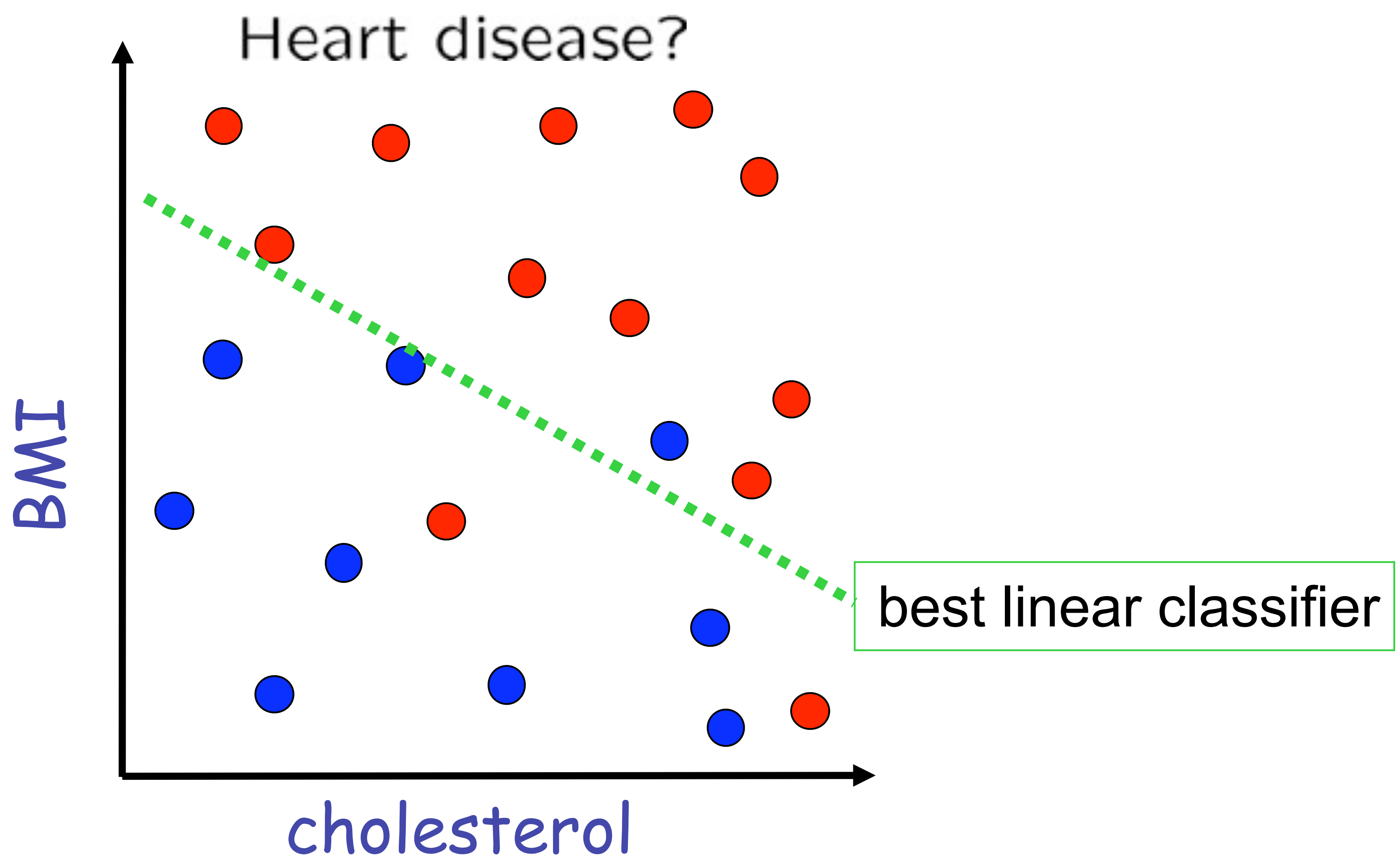
Active Learning

Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^n$



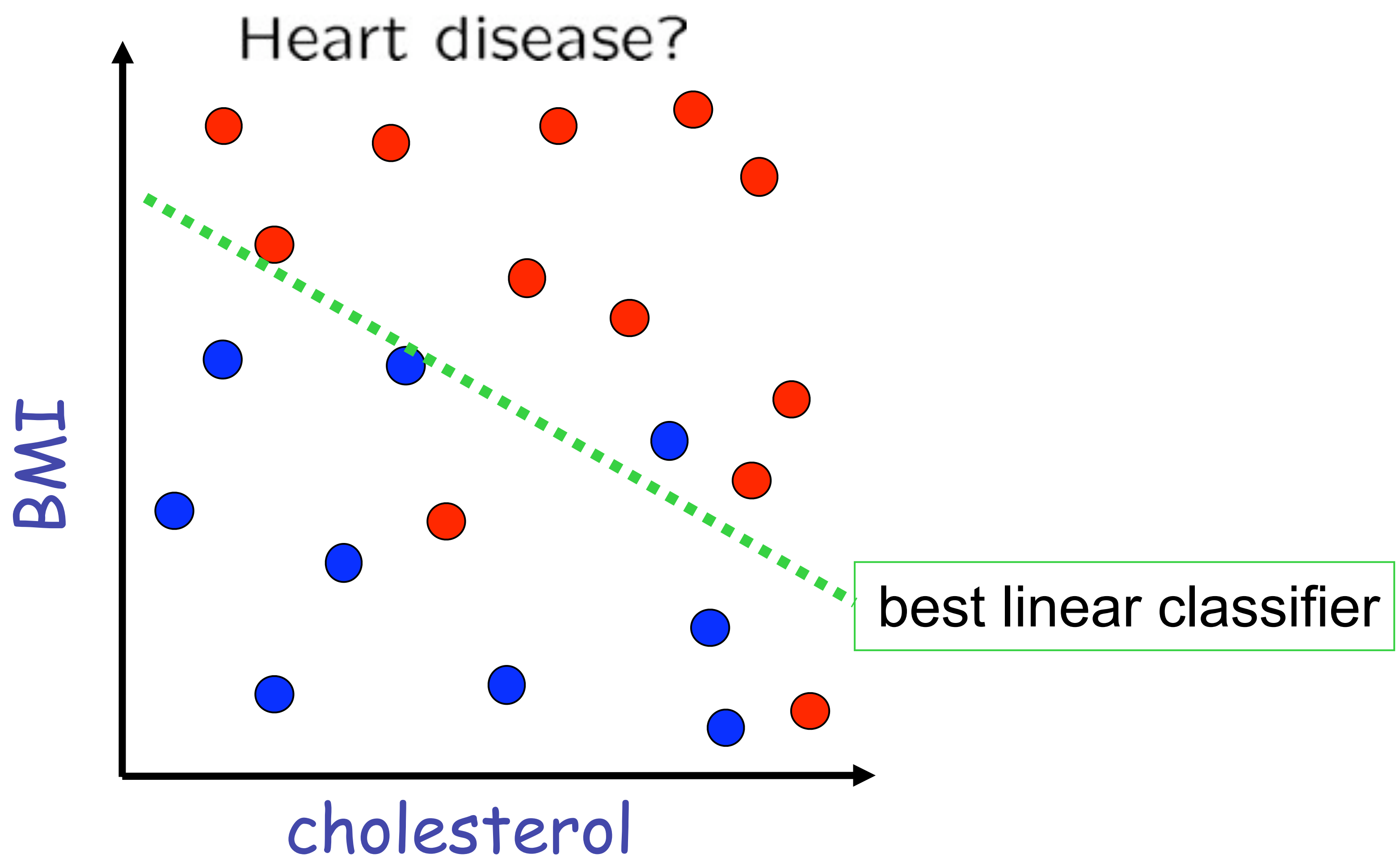
Active Learning

Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^n$



Active Learning

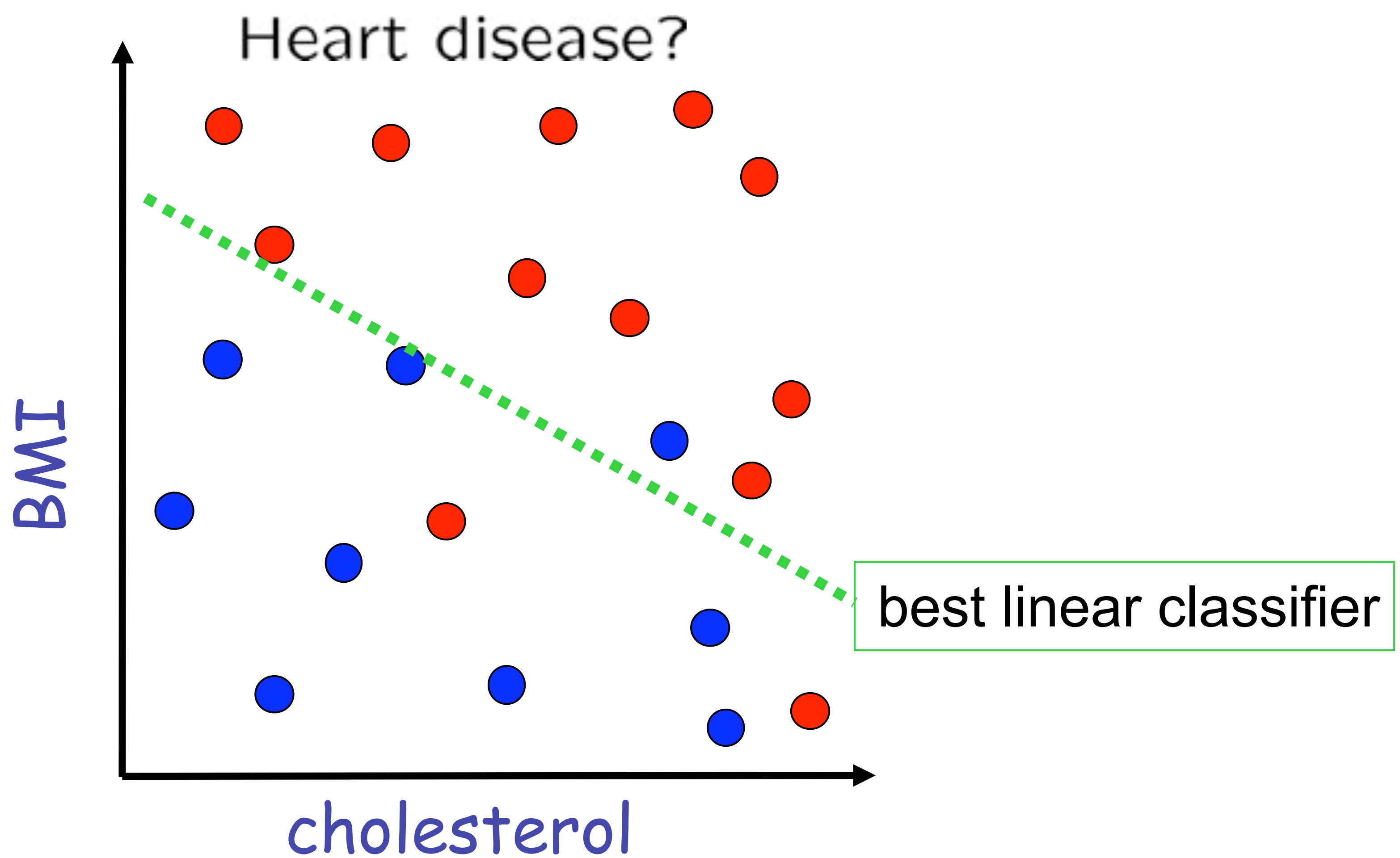
Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^n$



Passive Learning: training examples selected at random

Active Learning

Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^n$

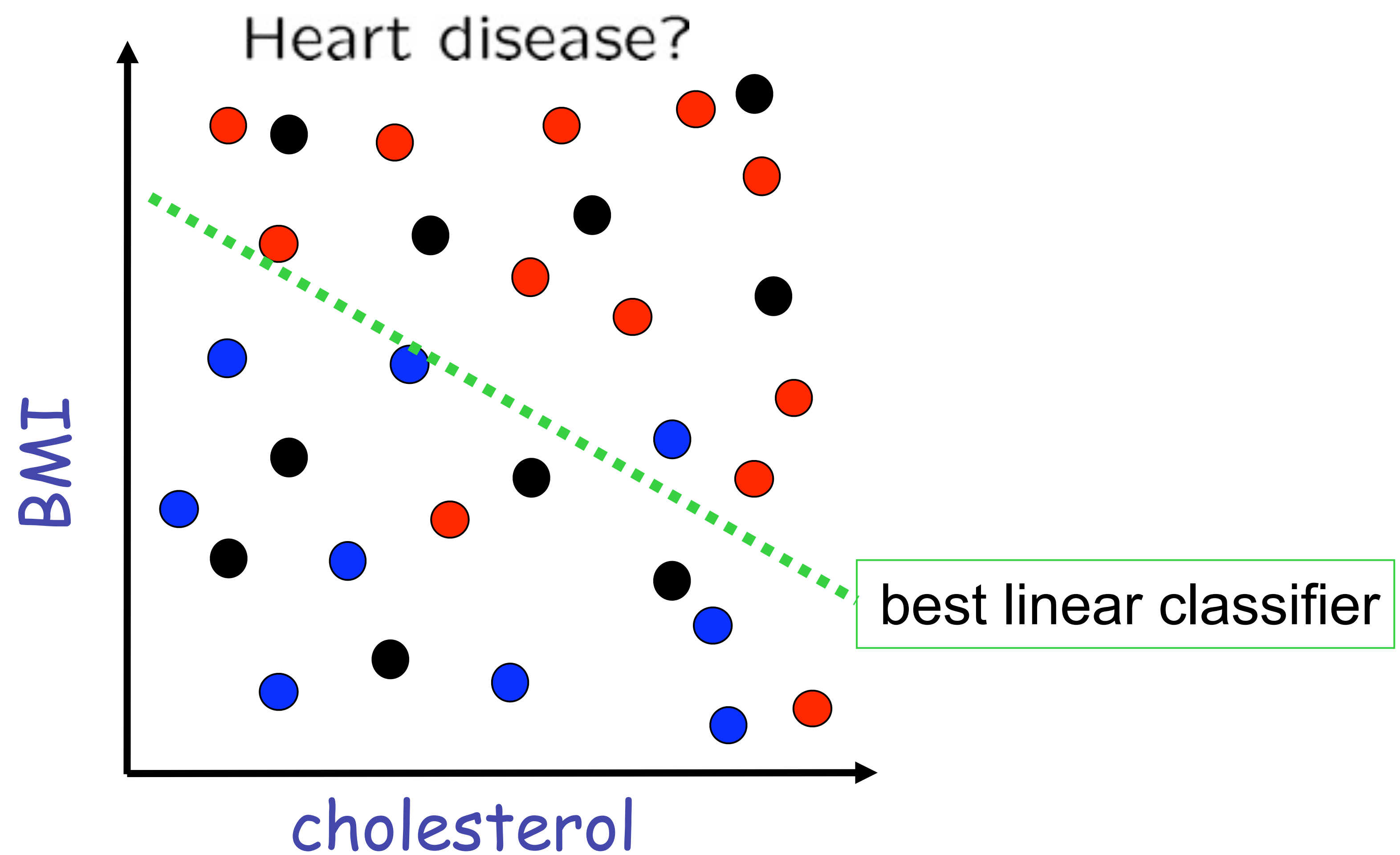


Passive Learning: training examples selected at random

Active Learning: especially informative examples are sequentially selected

Active Learning

Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^n$

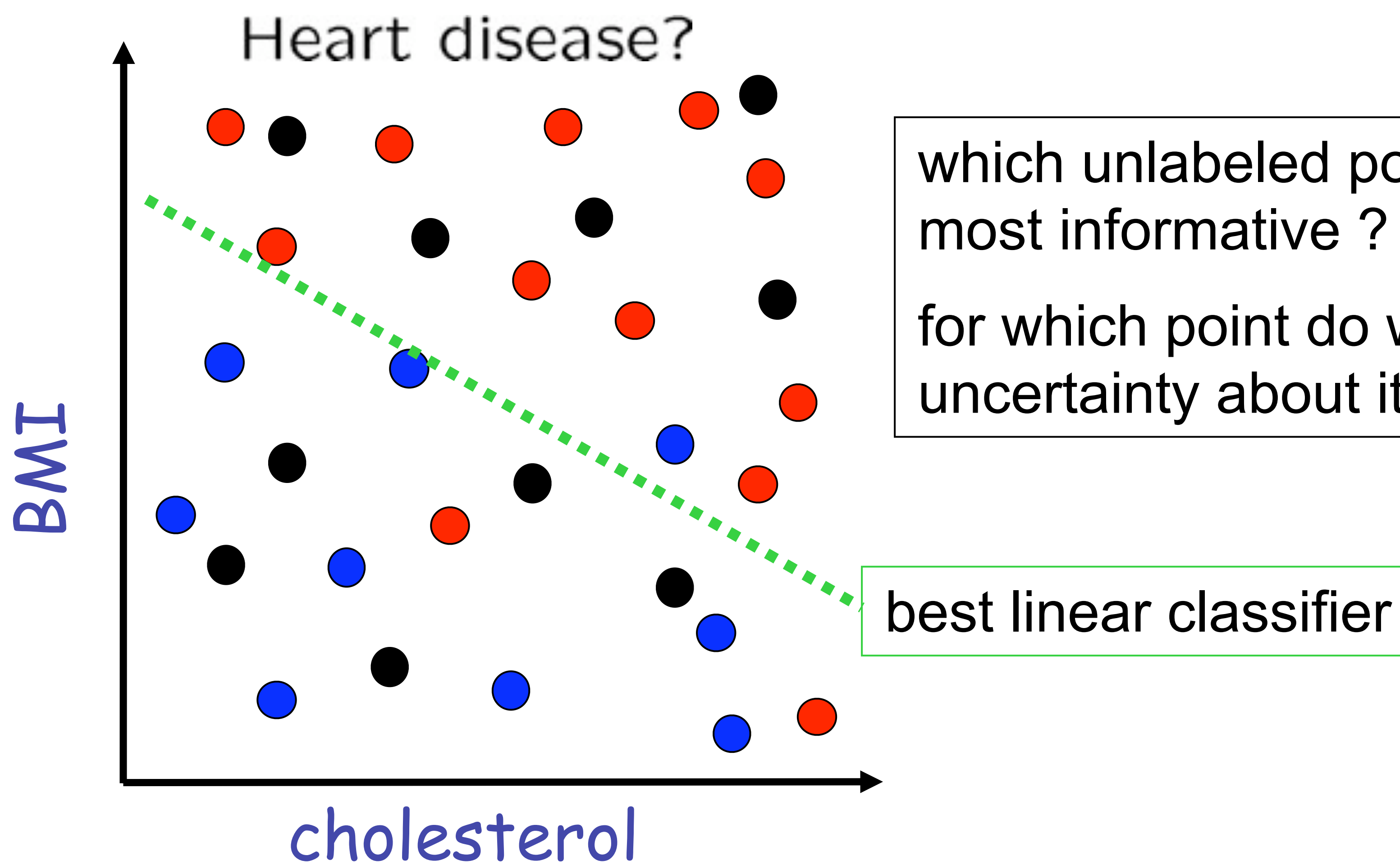


Passive Learning: training examples selected at random

Active Learning: especially informative examples are sequentially selected

Active Learning

Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^{n}$



which unlabeled point is likely to be most informative ?

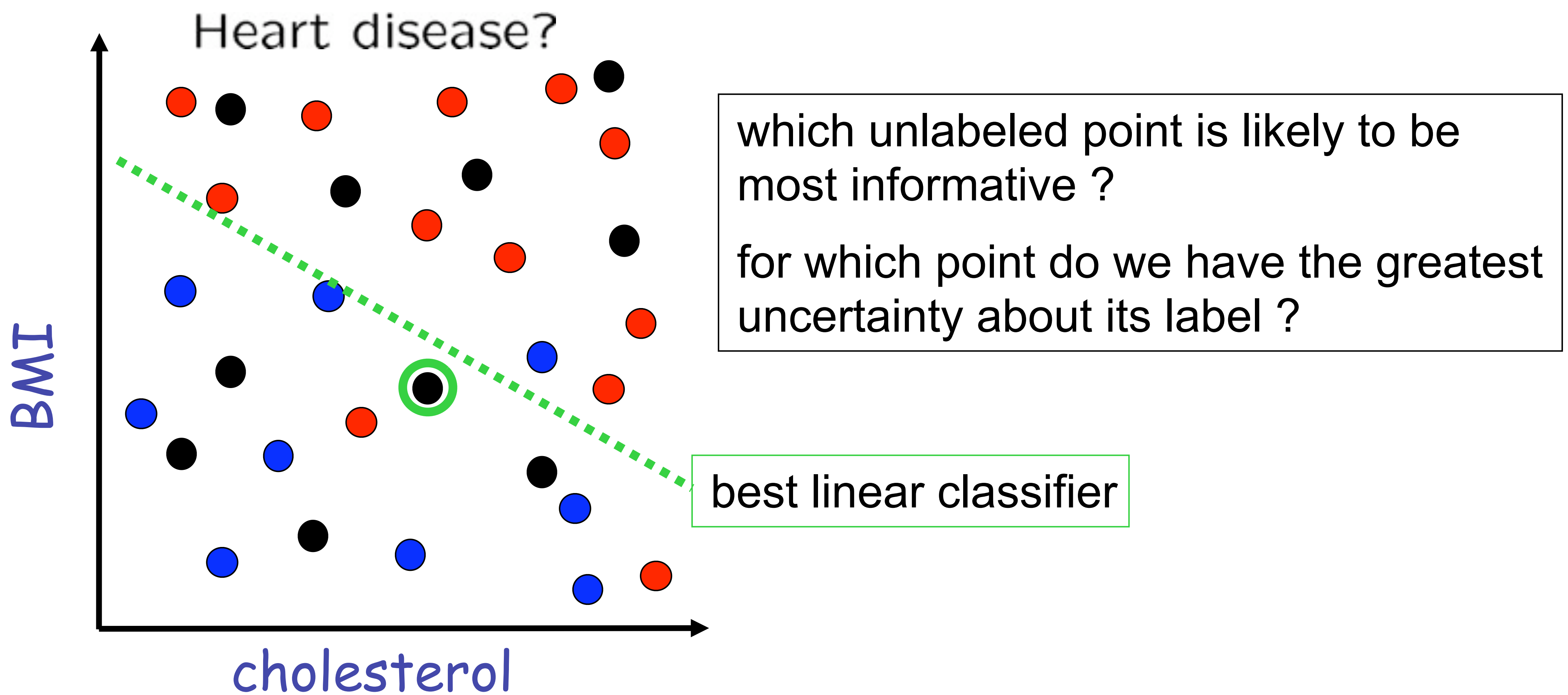
for which point do we have the greatest uncertainty about its label ?

Passive Learning: training examples selected at random

Active Learning: especially informative examples are sequentially selected

Active Learning

Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^n$

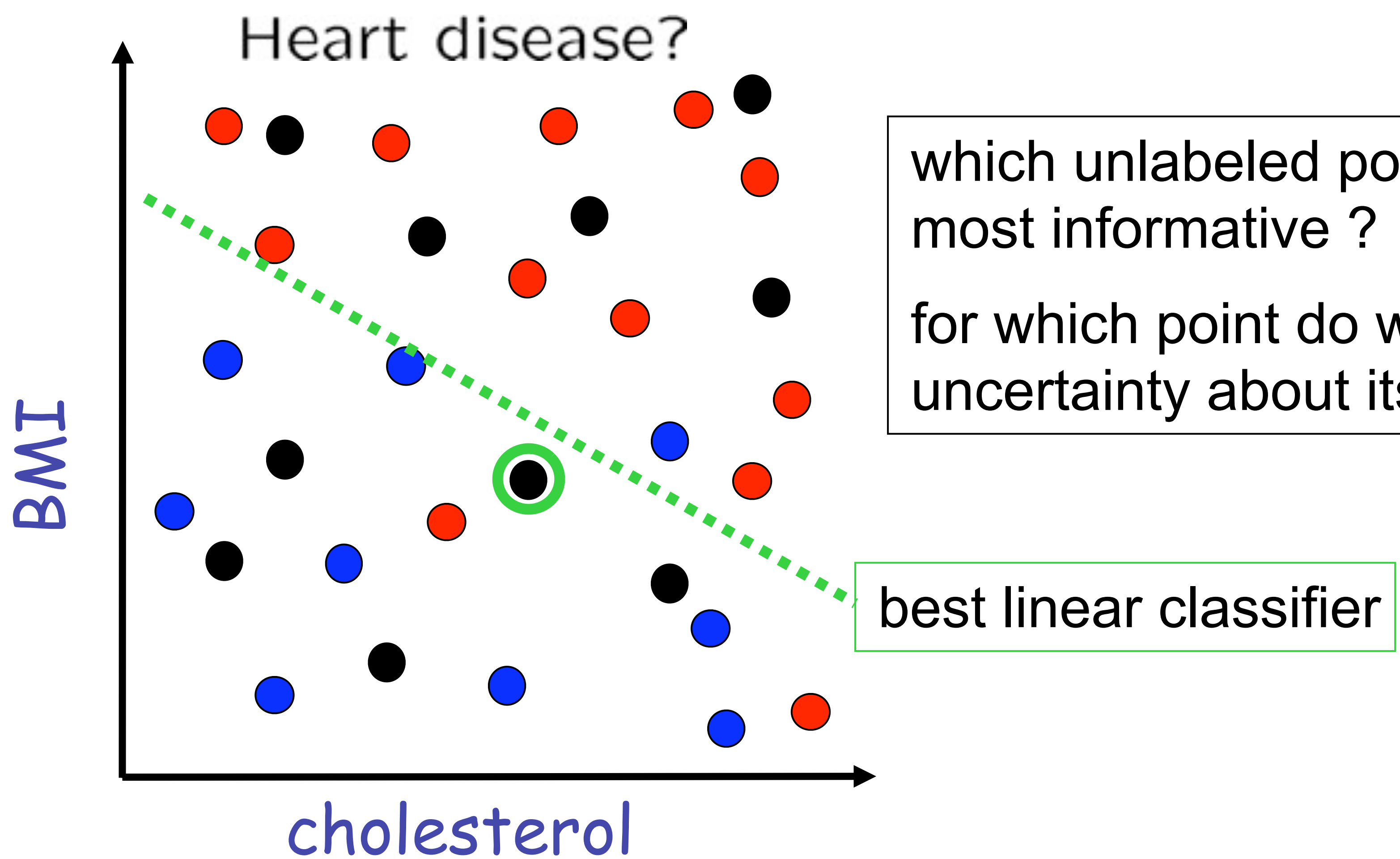


Passive Learning: training examples selected at random

Active Learning: especially informative examples are sequentially selected

Active Learning

Learn to predict labels y from features x based on training examples $\{(x_i, y_i)\}_{i=1}^n$



which unlabeled point is likely to be most informative ?

for which point do we have the greatest uncertainty about its label ?

Passive Learning: training examples selected at random

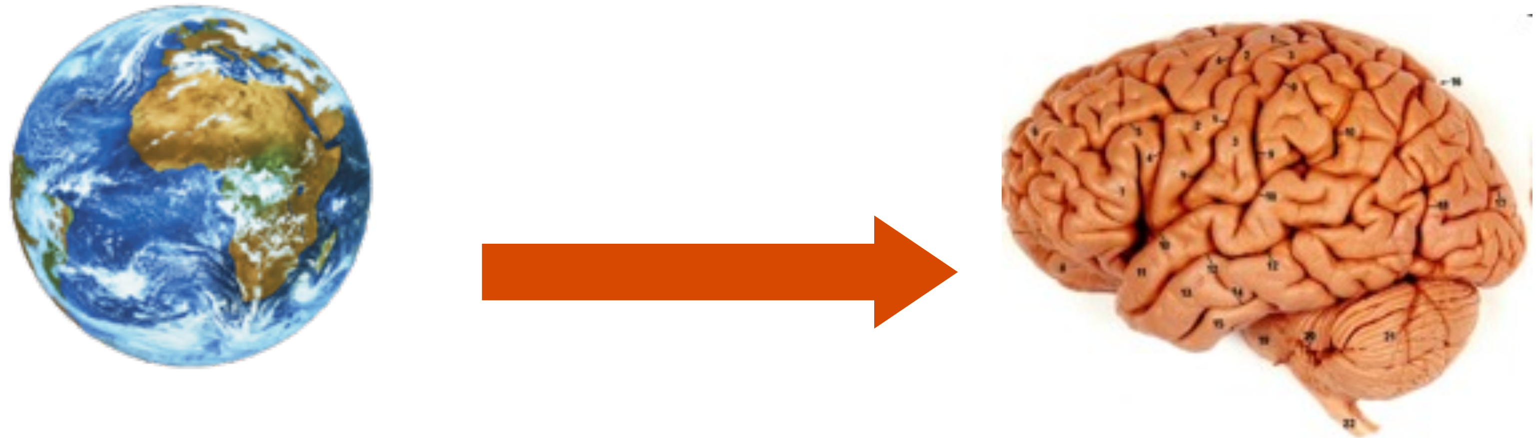
Active Learning: especially informative examples are sequentially selected

Active learning can very effectively "narrow down" the location of the optimal decision boundary

The Theory of the Organism-Environment System: III. Role of Efferent Influences on Receptors in the Formation of Knowledge*

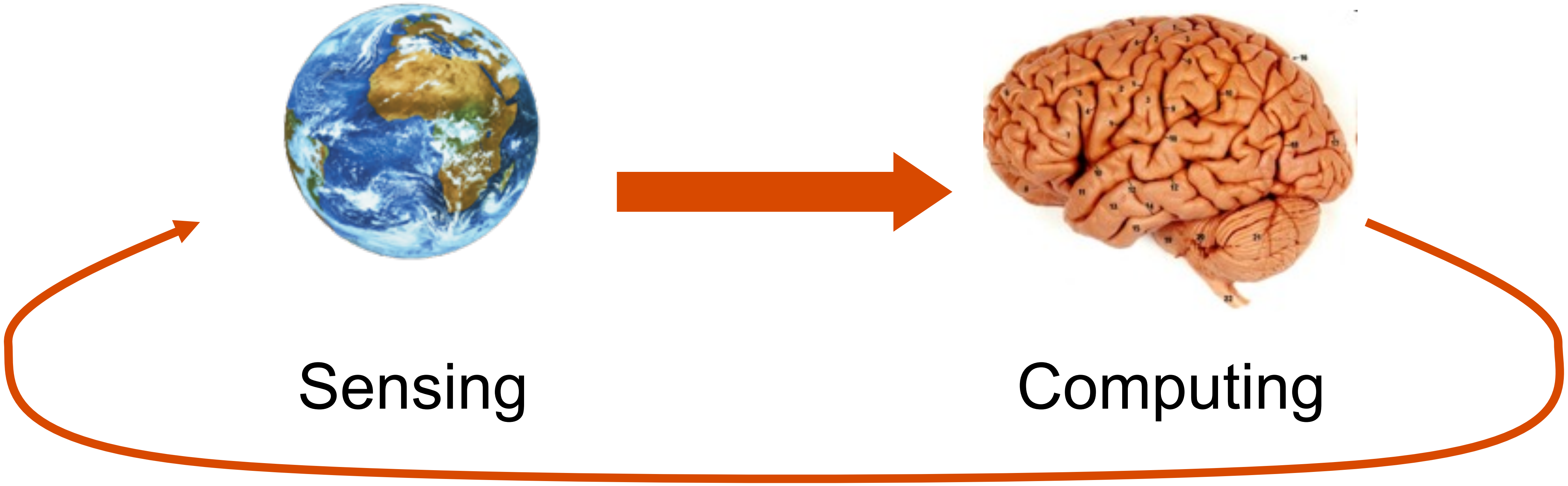
TIMO JARVILEHTO
Department of Behavioral Sciences, University of Oulu, Finland

Abstract—The present article is an attempt to give—in the frame of the theory of the organism-environment system (Jarvilehto, 1998a)—a new interpretation to the role of efferent influences on receptor activity and to the functions of senses in the formation of knowledge. It is argued, on the basis of experimental evidence and theoretical considerations, that the senses are not transmitters of environmental information, but create a direct connection between the organism and the environment, which makes the development of a dynamic living system, the organism-environment system, possible. In this connection process, the efferent influences on receptor activity are of particular significance because, with their help, the receptors may be adjusted in relation to the parts of the environment that are most important in achieving behavioral results. Perception is the process of joining of new parts of the environment to the organism-environment system; thus, the formation of knowledge by perception is based on reorganization (widening and differentiation) of the organism-environment system, and not on transmission of information from the environment. With the help of the efferent influences on receptors, each organism creates its own peculiar world that is simultaneously subjective and objective. The present considerations have far-reaching influences as well on experimental work in neurophysiology and psychology of perception as on philosophical considerations of knowledge formation.

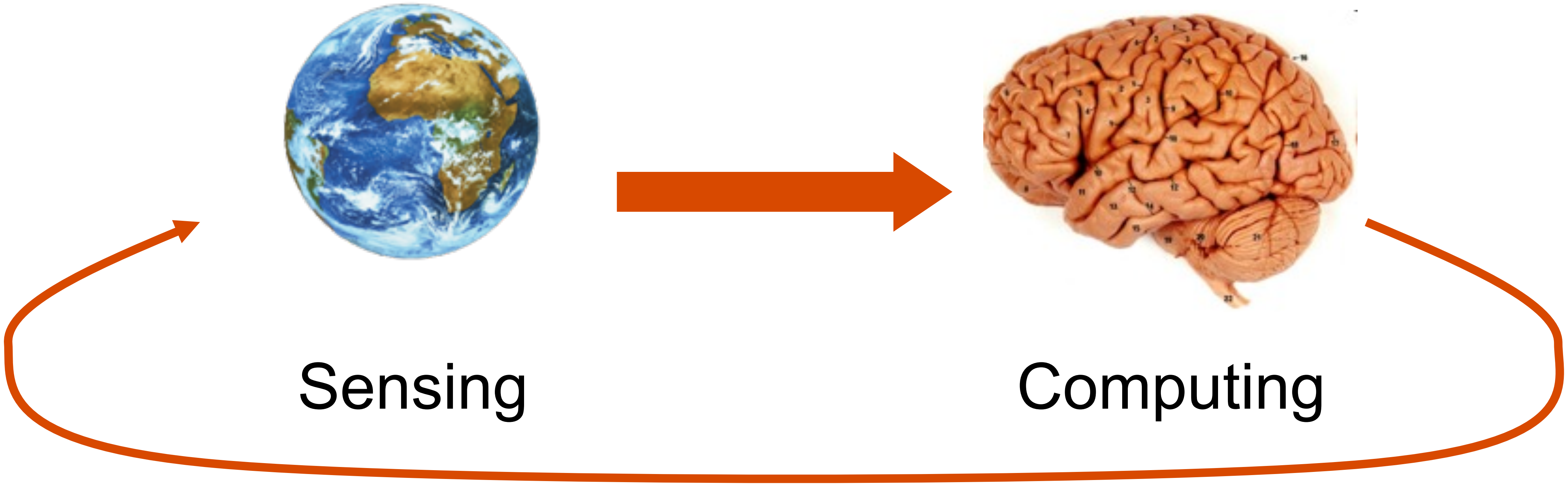


Sensing Computing

Abstract—The present article is an attempt to give—in the frame of the theory of the organism-environment system (Jarvilehto, 1998a)—a new interpretation to the role of efferent influences on receptor activity and to the functions of senses in the formation of knowledge. It is argued, on the basis of experimental evidence and theoretical considerations, that the senses are not transmitters of environmental information, but create a direct connection between the organism and the environment, which makes the development of a dynamic living system, the organism-environment system, possible. In this connection process, the efferent influences on receptor activity are of particular significance because, with their help, the receptors may be adjusted in relation to the parts of the environment that are most important in achieving behavioral results. Perception is the process of joining of new parts of the environment to the organism-environment system; thus, the formation of knowledge by perception is based on reorganization (widening and differentiation) of the organism-environment system, and not on transmission of information from the environment. With the help of the efferent influences on receptors, each organism creates its own peculiar world that is simultaneously subjective and objective. The present considerations have far-reaching influences as well on experimental work in neurophysiology and psychology of perception as on philosophical considerations of knowledge formation.



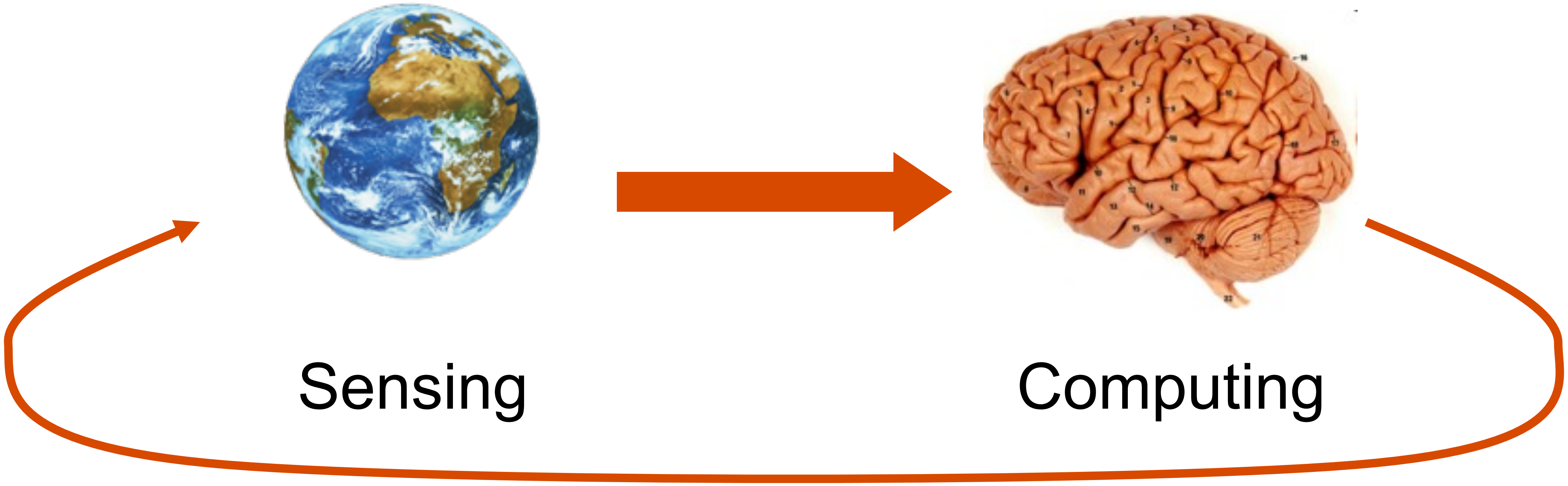
Abstract—The present article is an attempt to give—in the frame of the theory of the organism-environment system (Jarvilehto, 1998a)—a new interpretation to the role of efferent influences on receptor activity and to the functions of senses in the formation of knowledge. It is argued, on the basis of experimental evidence and theoretical considerations, that the senses are not transmitters of environmental information, but create a direct connection between the organism and the environment, which makes the development of a dynamic living system, the organism-environment system, possible. In this connection process, the efferent influences on receptor activity are of particular significance because, with their help, the receptors may be adjusted in relation to the parts of the environment that are most important in achieving behavioral results. Perception is the process of joining of new parts of the environment to the organism-environment system; thus, the formation of knowledge by perception is based on reorganization (widening and differentiation) of the organism-environment system, and not on transmission of information from the environment. With the help of the efferent influences on receptors, each organism creates its own peculiar world that is simultaneously subjective and objective. The present considerations have far-reaching influences as well on experimental work in neurophysiology and psychology of perception as on philosophical considerations of knowledge formation.



Abstract—The present article is an attempt to give—in the frame of the theory of the organism-environment system (Jarvilehto, 1998a)—a new interpretation to the role of efferent influences on receptor activity and to the functions of senses in the formation of knowledge. It is argued, on the basis of experimental evidence and theoretical considerations, that the senses are not transmitters of environmental information, but create a direct

the senses are not transmitters of environmental information, but create a direct connection between the organism and the environment, that makes the develop of

process, the efferent influences on receptor activity are of particular significance because, with their help, the receptors may be adjusted in relation to the parts of the environment that are most important in achieving behavioral results. Perception is the process of joining of new parts of the environment to the organism-environment system; thus, the formation of knowledge by perception is based on reorganization (widening and differentiation) of the organism-environment system, and not on transmission of information from the environment. With the help of the efferent influences on receptors, each organism creates its own peculiar world that is simultaneously subjective and objective. The present considerations have far-reaching influences as well on experimental work in neurophysiology and psychology of perception as on philosophical considerations of knowledge formation.



Abstract—The present article is an attempt to give—in the frame of the theory of the organism-environment system (Jarvilehto, 1998a)—a new interpretation to the role of efferent influences on receptor activity and to the functions of senses in the formation of knowledge. It is argued, on the basis of experimental evidence and theoretical considerations, that the senses are not transmitters of environmental information, but create a direct

the senses are not transmitters of environmental information, but create a direct connection between the organism and the environment, that makes the develop of

process, the efferent influences on receptor activity are of particular significance because,

the receptors may be adjusted in relation to the parts of the environment that are most important in achieving behavioral results. Perception is the process of joining

of new parts of the environment to the organism-environment system; thus, the formation of knowledge by perception is based on reorganization (widening and differentiation) of the organism-environment system, and not on transmission of information from the environment. With the help of the efferent influences on receptors, each organism creates its own peculiar world that is simultaneously subjective and objective. The present considerations have far-reaching influences as well on experimental work in neurophysiology and psychology of perception as on philosophical considerations of knowledge formation.

Visual Perception

Attentional mechanisms probably limit our capacity to about 44 bits per-glimpse (Verghese and Pelli (1992))

So how to we perceive 'reality' from so few bits of information?

Visual Perception

Attentional mechanisms probably limit our capacity to about 44 bits per-glimpse (Verghese and Pelli (1992))

So how to we perceive ‘reality’ from so few bits of information?

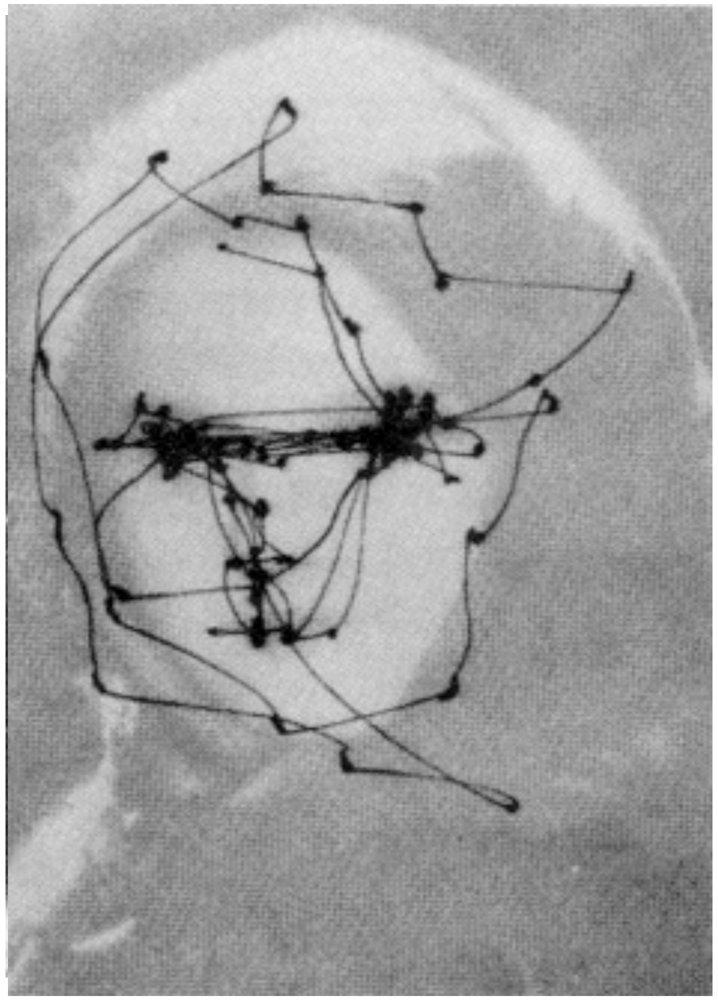
Churchland, Ramachandran, & Sejnowksi ’94: **“Interactive vision** is exploratory and predictive. Visual learning allows an animal to predict what will happen in the future; behavior, such as eye movements, aids in updating and upgrading the predictive representations.”

Visual Perception

Attentional mechanisms probably limit our capacity to about 44 bits per-glimpse (Verghese and Pelli (1992))

So how to we perceive 'reality' from so few bits of information?

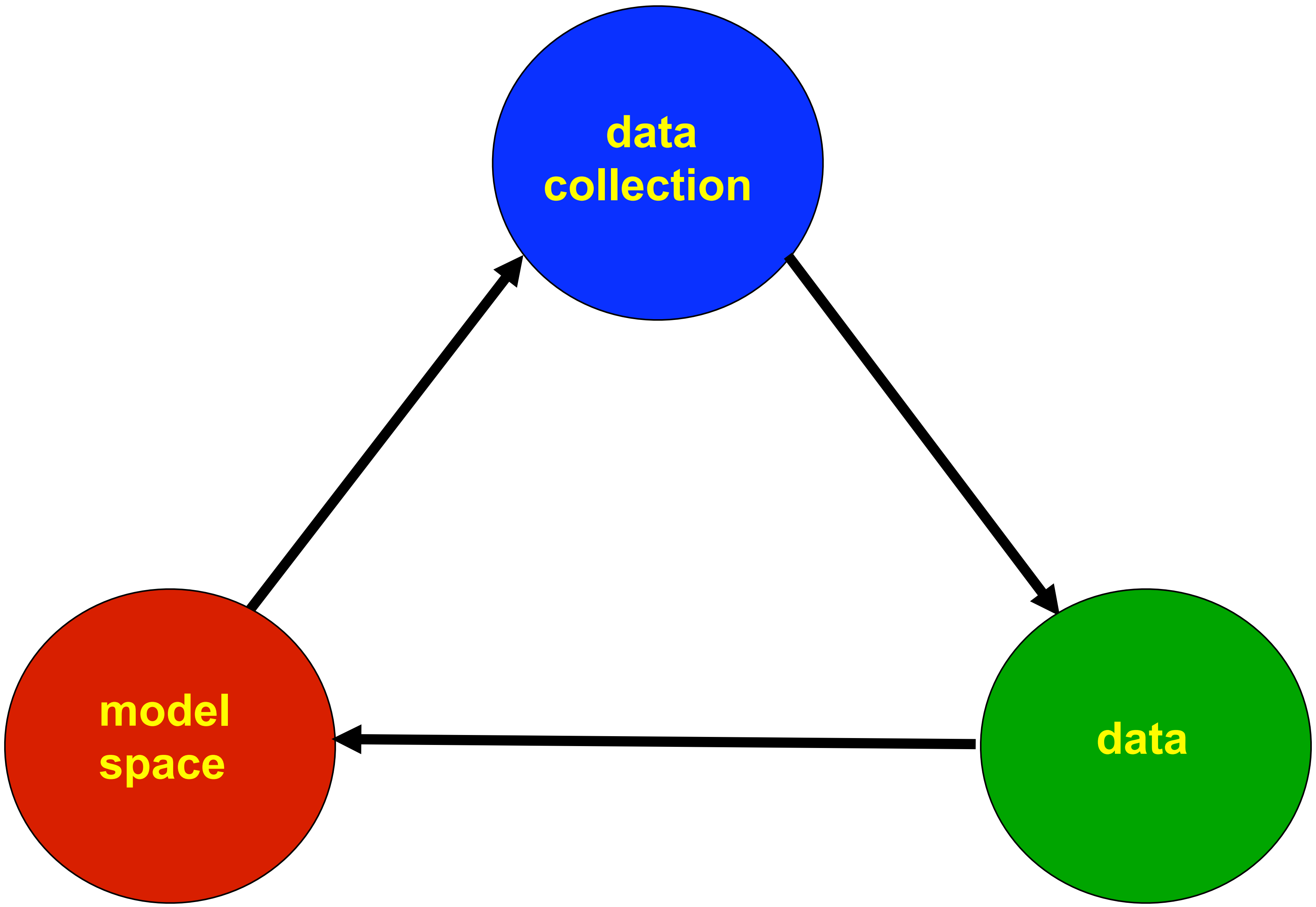
Churchland, Ramachandran, & Sejnowksi '94: "**Interactive vision** is exploratory and predictive. Visual learning allows an animal to predict what will happen in the future; behavior, such as eye movements, aids in updating and upgrading the predictive representations."




S

Mathematical Theory of Active Sensing and Learning

$\mathcal{Y}$: possible measurements/experiments



$\mathcal{X}$: models/hypotheses
under consideration

$y_1(x), y_2(x), \ldots$: information/data

Adaptive vs. Non-Adaptive: Three Situations

The "bare minimum" number of measurements depends on intrinsic complexity of $\mathcal{X}$. In practice, the minimum number depends on jointly on $\mathcal{X}$ and $\mathcal{Y}$.

Adaptive vs. Non-Adaptive: Three Situations

The "bare minimum" number of measurements depends on intrinsic complexity of $\mathcal{X}$. In practice, the minimum number depends on jointly on $\mathcal{X}$ and $\mathcal{Y}$.

Equal and Good:
adaptive and non-adaptive equally informative and require about the bare minimum of measurements

Adaptive vs. Non-Adaptive: Three Situations

The "bare minimum" number of measurements depends on intrinsic complexity of $\mathcal{X}$. In practice, the minimum number depends on jointly on $\mathcal{X}$ and $\mathcal{Y}$.

Equal and Good:
adaptive and non-adaptive equally informative and require about the bare minimum of measurements

Equal and Bad:
adaptive and non-adaptive equally (non)-informative and require many more measurements then the bare minimum

Adaptive vs. Non-Adaptive: Three Situations

The "bare minimum" number of measurements depends on intrinsic complexity of $\mathcal{X}$. In practice, the minimum number depends on jointly on $\mathcal{X}$ and $\mathcal{Y}$.

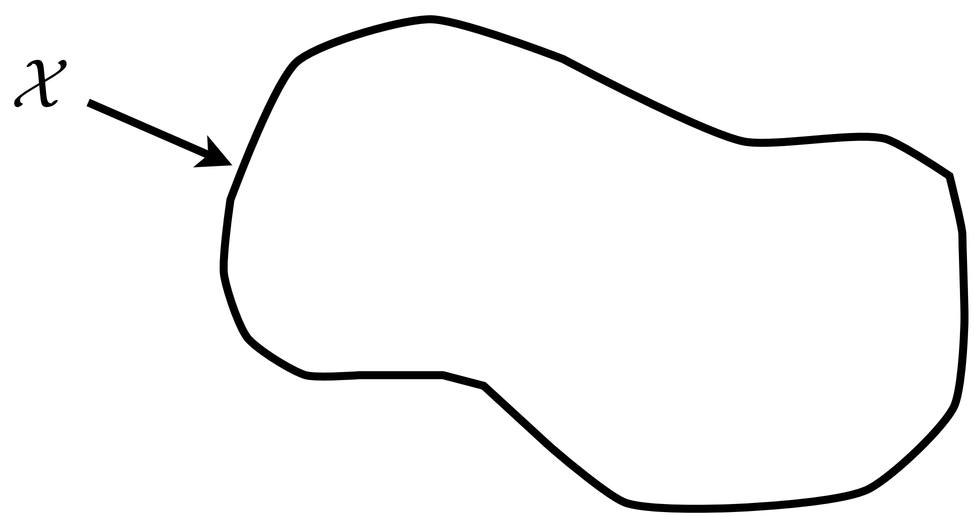
Equal and Good: adaptive and non-adaptive equally informative and require about the bare minimum of measurements

Equal and Bad: adaptive and non-adaptive equally (non)-informative and require many more measurements then the bare minimum

Good and Bad: adaptive requires bare minimum number of measurments, non-adaptive requires many more

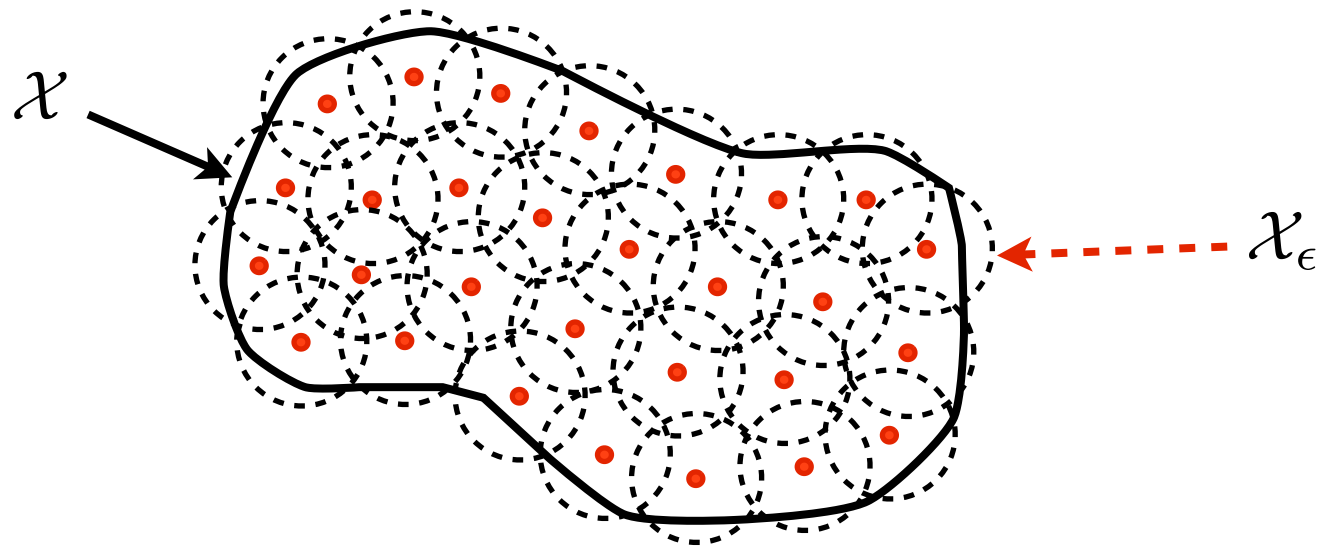
The Bare Minimum

Assume $\mathcal{X}$ is equipped with metric d and is compact.



The Bare Minimum

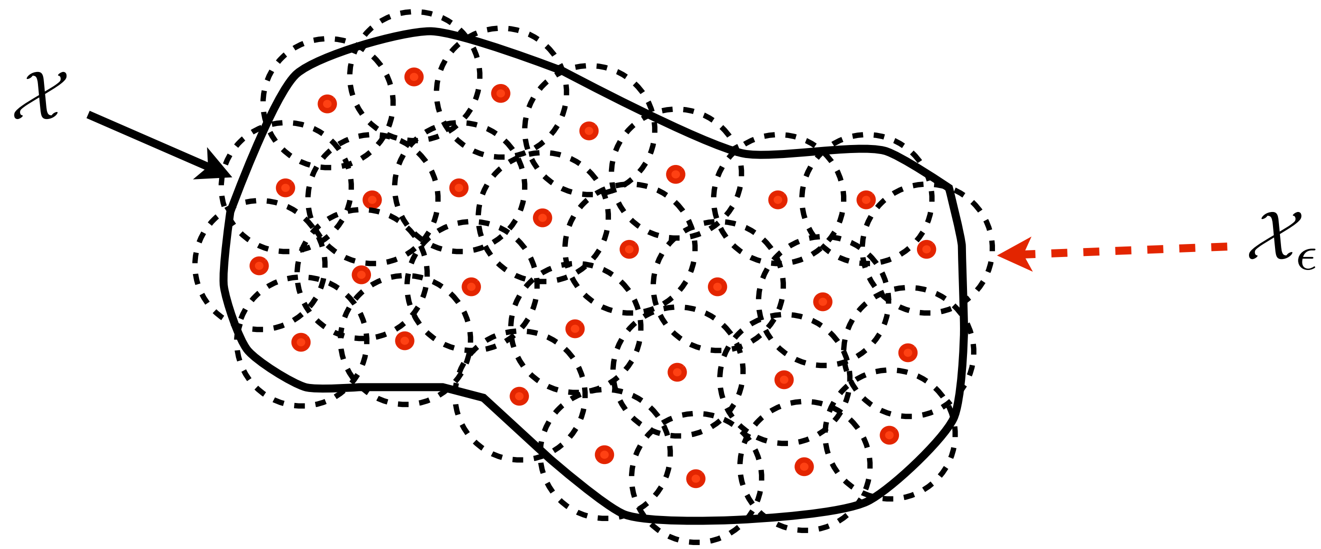
Assume $\mathcal{X}$ is equipped with metric d and is compact.



Let $\mathcal{X}_\epsilon \subset \mathcal{X}$ be a finite subset of size N_ϵ having the property that any element of $\mathcal{X}$ is within distance ϵ of an element in $\mathcal{X}_\epsilon$

The Bare Minimum

Assume $\mathcal{X}$ is equipped with metric d and is compact.

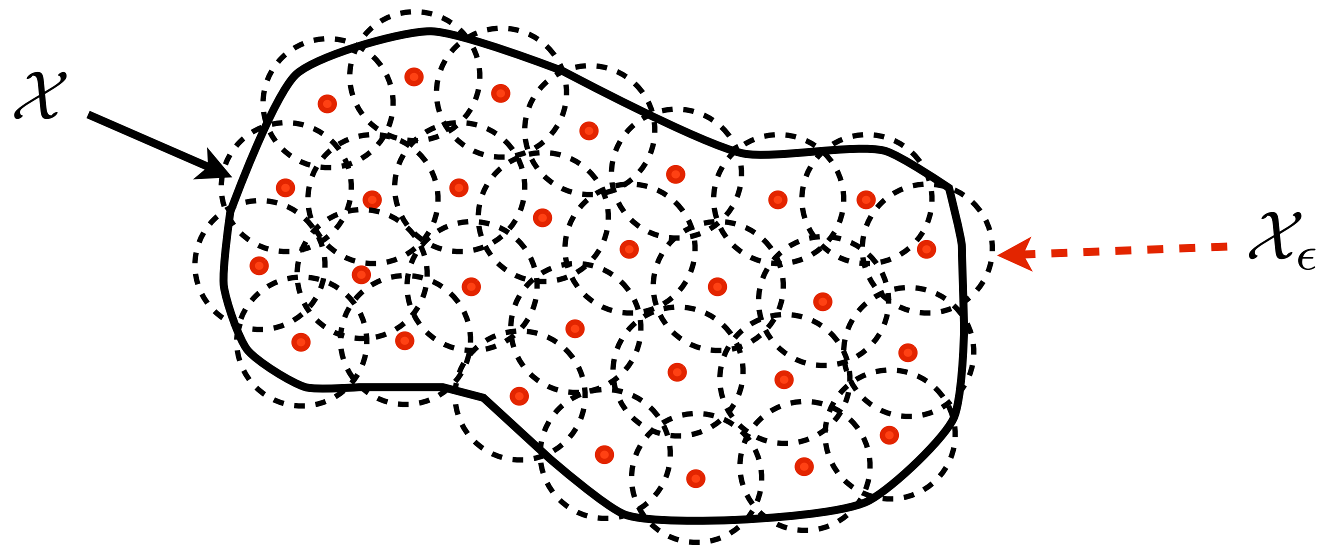


Let $\mathcal{X}_\epsilon \subset \mathcal{X}$ be a finite subset of size N_ϵ having the property that any element of $\mathcal{X}$ is within distance ϵ of an element in $\mathcal{X}_\epsilon$

Metric Entropy: Need at least $\log N_\epsilon$ bits of information to approximately determine any $x \in \mathcal{X}$

The Bare Minimum

Assume $\mathcal{X}$ is equipped with metric d and is compact.



Let $\mathcal{X}_\epsilon \subset \mathcal{X}$ be a finite subset of size N_ϵ having the property that any element of $\mathcal{X}$ is within distance ϵ of an element in $\mathcal{X}_\epsilon$

Metric Entropy: Need at least $\log N_\epsilon$ bits of information to approximately determine any $x \in \mathcal{X}$

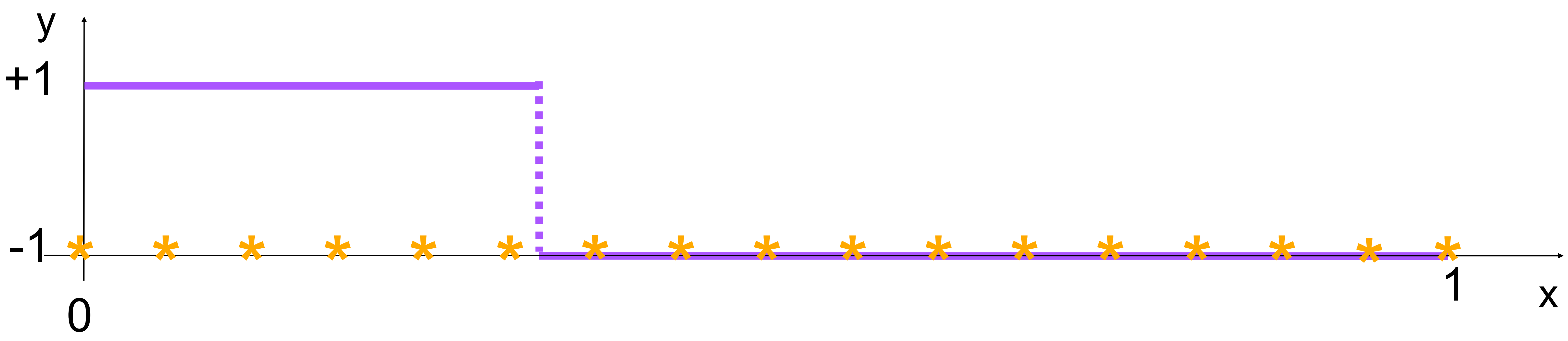
Ex. suppose $\mathcal{X} = [0,1]^d$. we can take a uniform grid of points spaced ϵ apart as our cover. Then $N_\epsilon = (\frac{1}{\epsilon})^d$ and $\log N_\epsilon = d \log(1/\epsilon)$.

Binary Search

$$\mathcal{X} = \{\,\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0, 1]\,\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

Binary Search

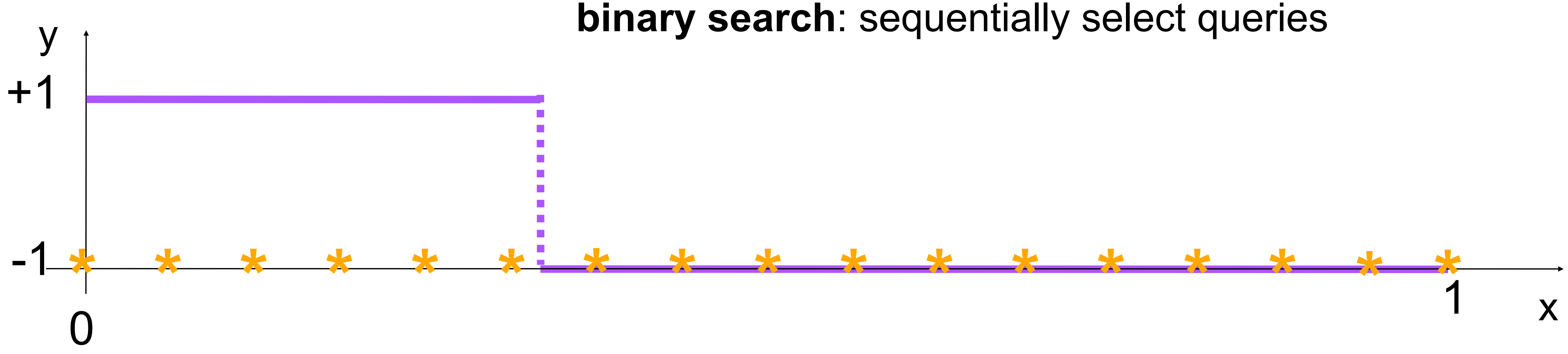
$$\mathcal{X} = \{ \text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0, 1] \}$$
$$\mathcal{Y} = \text{“membership queries”}$$



Binary Search

$$\mathcal{X} = \{\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0, 1]\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

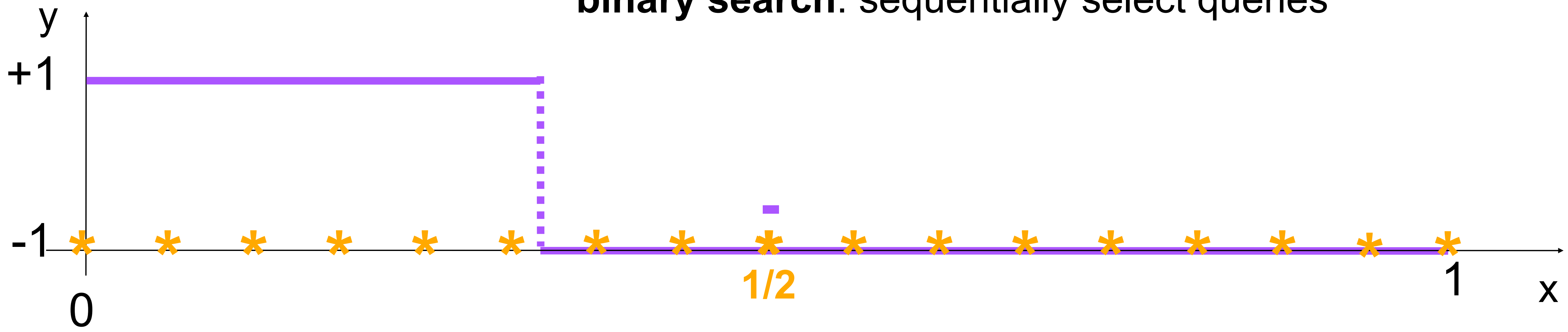
binary search: sequentially select queries



Binary Search

$$\mathcal{X} = \{\,\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0, 1]\,\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

binary search: sequentially select queries

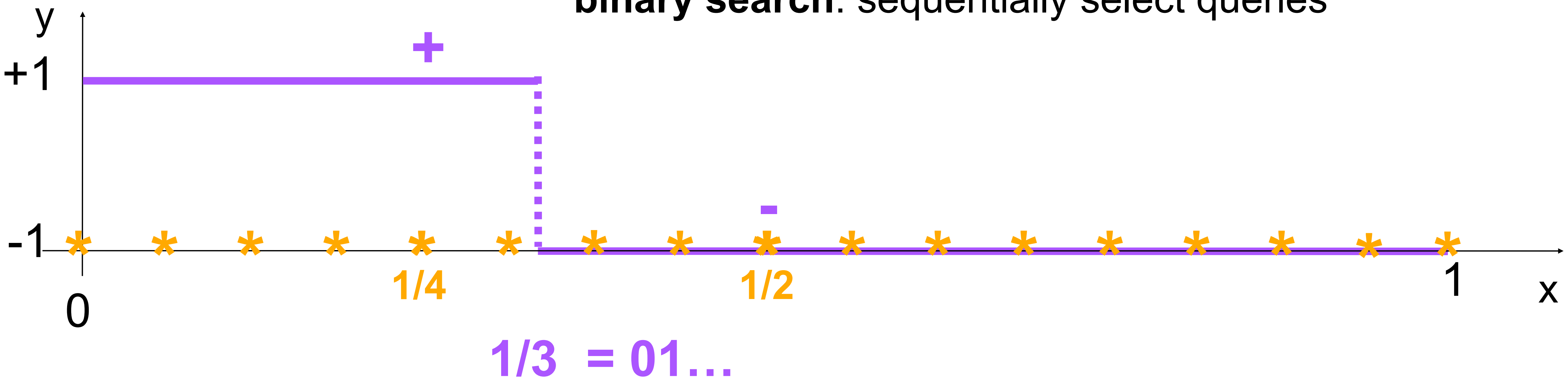


1/3 = 0…

Binary Search

$$\mathcal{X} = \{ \text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0, 1] \}$$
$$\mathcal{Y} = \text{“membership queries”}$$

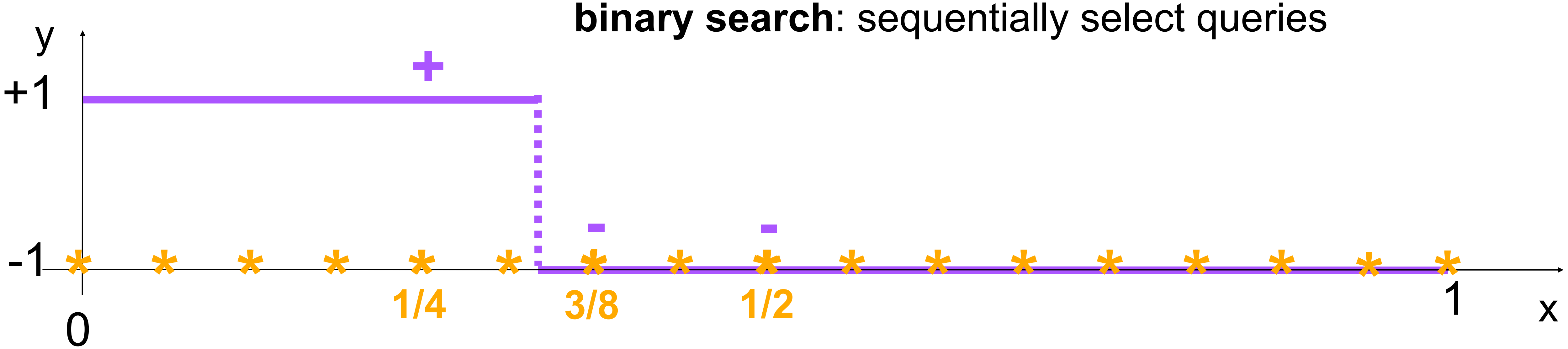
binary search: sequentially select queries



Binary Search

$$\mathcal{X} = \{\,\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0,1]\,\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

binary search: sequentially select queries

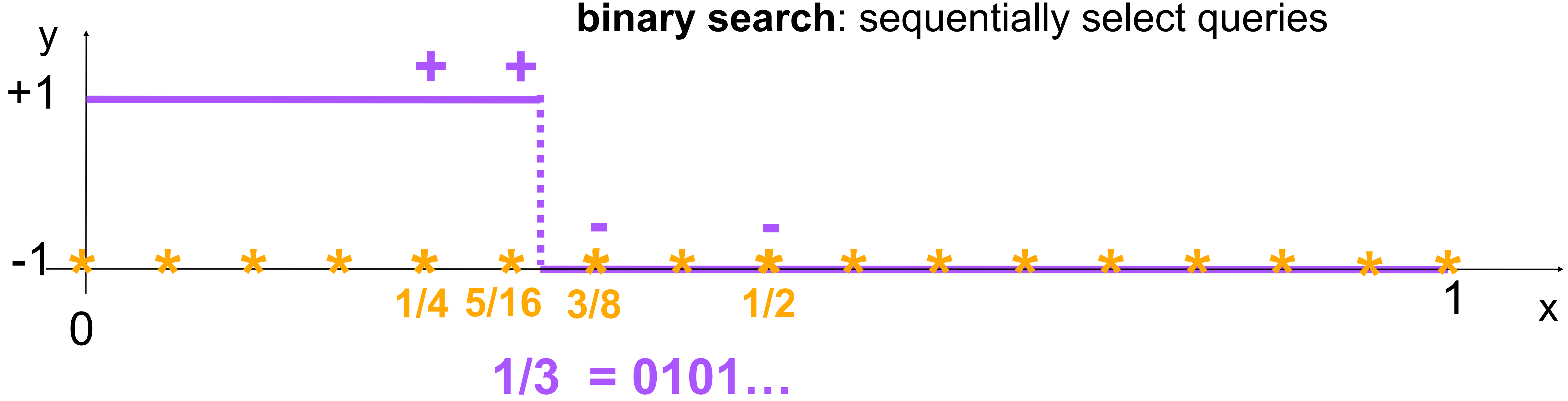


1/3 = 010…

Binary Search

$$\mathcal{X} = \{\,\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0,1]\,\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

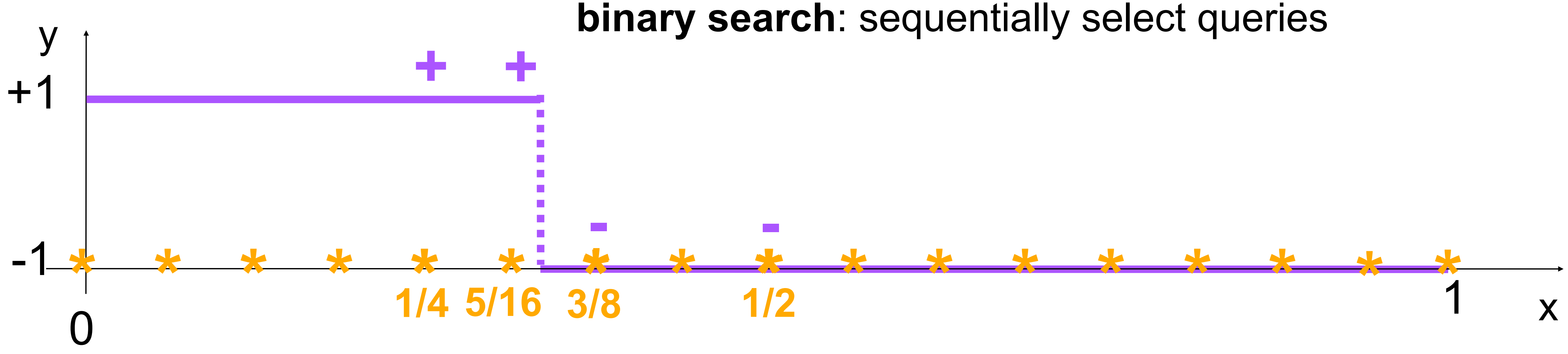
binary search: sequentially select queries



Binary Search

$$\mathcal{X} = \{\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0,1]\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

binary search: sequentially select queries



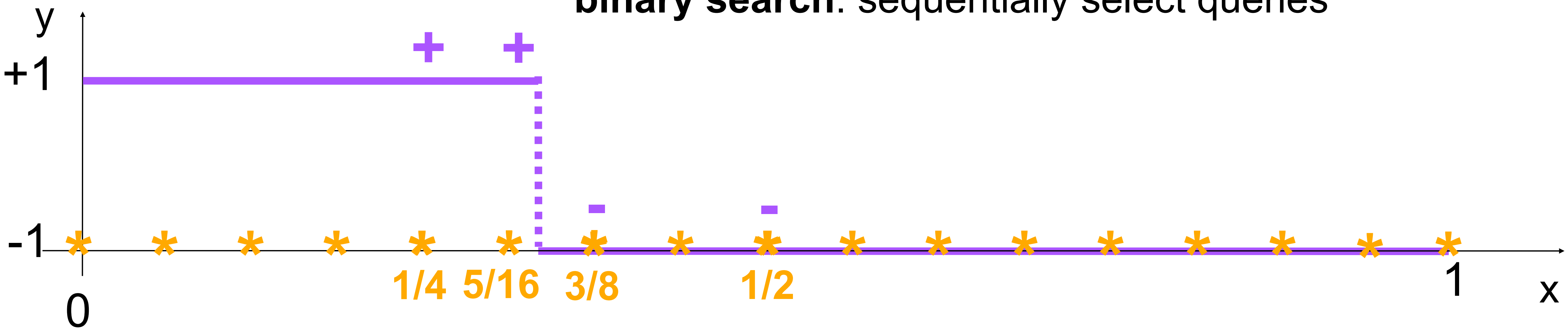
1/3 = 0101…

requires $\log_2 N$ queries

Binary Search

$$\mathcal{X} = \{\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0,1]\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

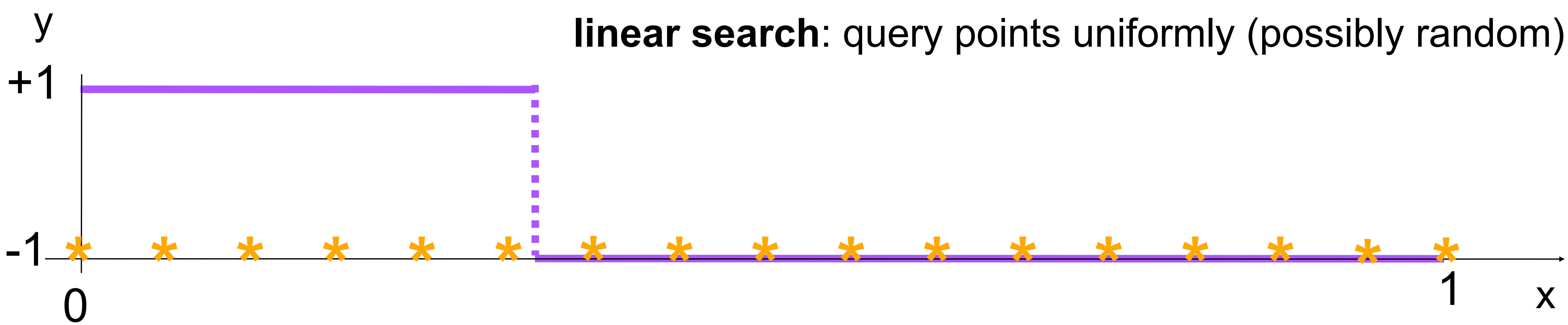
binary search: sequentially select queries



1/3 = 0101…

requires $\log_2 N$ queries

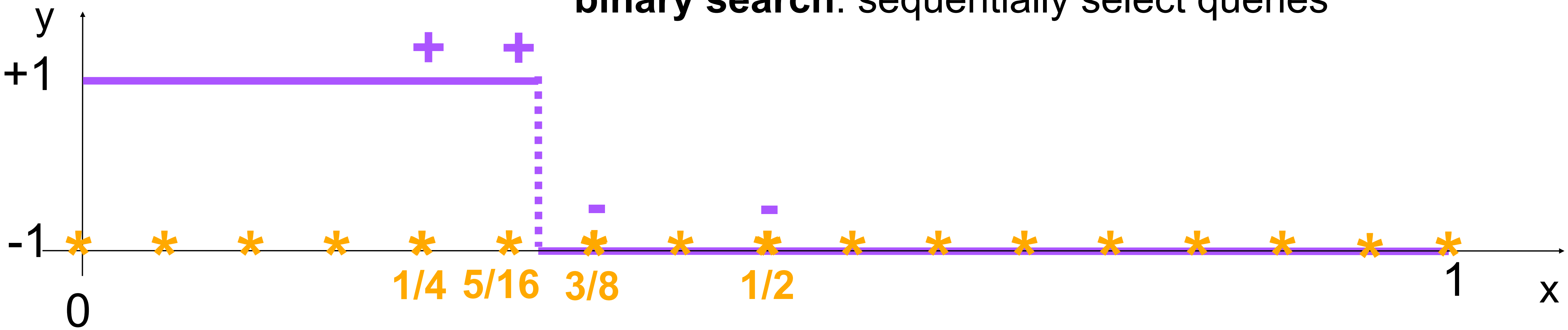
linear search: query points uniformly (possibly random)



Binary Search

$$\mathcal{X} = \{\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0,1]\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

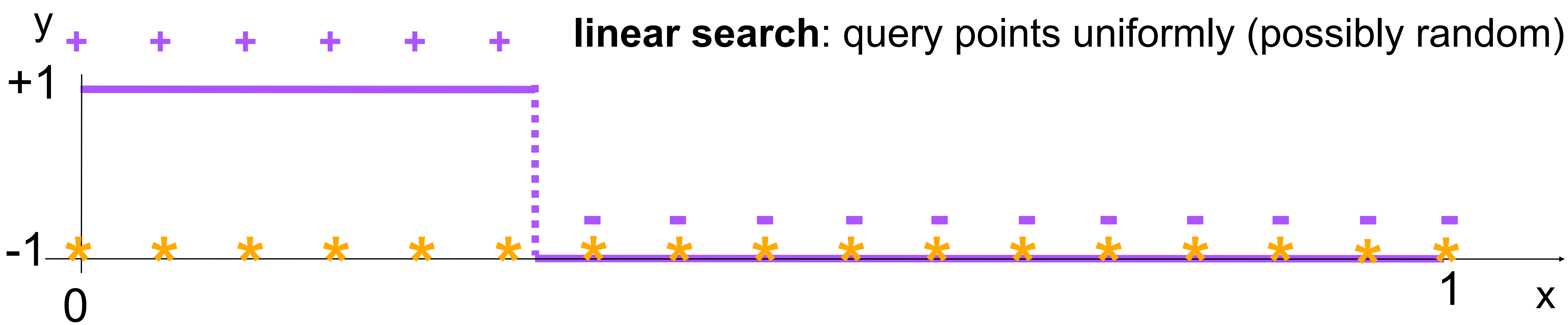
binary search: sequentially select queries



1/3 = 0101…

requires $\log_2 N$ queries

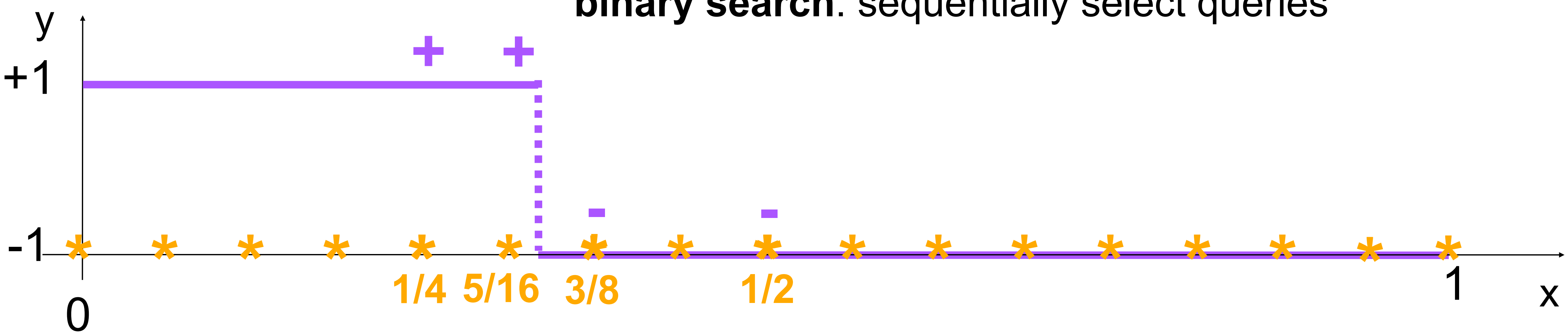
linear search: query points uniformly (possibly random)



Binary Search

$$\mathcal{X} = \{\,\text{subsets } [0, \tfrac{1}{N}], [0, \tfrac{2}{N}], \ldots, [0, 1]\,\}$$
$$\mathcal{Y} = \text{“membership queries”}$$

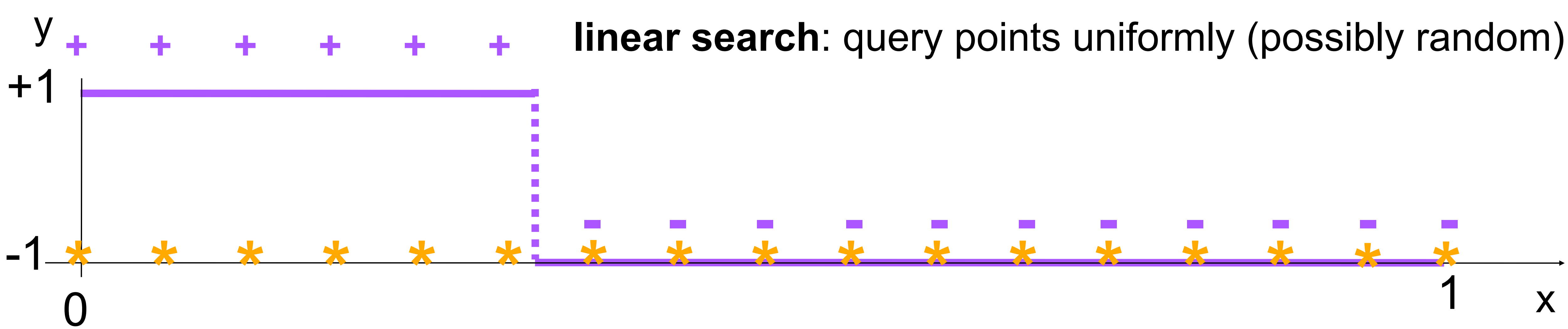
binary search: sequentially select queries



1/3 = 0101…

requires $\log_2 N$ queries

linear search: query points uniformly (possibly random)



requires $O(N)$ queries

Outline of Tutorial

Part 1: Introduction (Rob), 9:00-9:30

Part 2: Active Sensing (Jarvis), 9:30-10:30

Break, 10:30-10:45

Part 3: Active Learning (Rob), 10:45-11:45

Part 4: Conclusions and Future Directions, 11:45-12