# Measures of Structural Complexity

Complexity Lecture 2: Computational Mechanics (CSSS 2011); Jim Crutchfield

# Measures of Complexity ...

## Measures of Structural Complexity:

| Measures | | Interpretation |
|---|---|---|
| Entropy Rate | $h_\mu$ | Intrinsic Randomness |
| Excess Entropy | $\mathbf{E}$ | Info: Past to Future |
| Predictability Gain | $\mathbf{G}$ | Redundancy |
| Transient Information | $\mathbf{T}$ | Synchronization |

How related to statistical complexity?

How to get from $\epsilon\mathrm{M}$?

## Measures from the $\epsilon\mathrm{M}$:

### Entropy Rate of a Process:

$$h_\mu(\Pr(\overleftrightarrow{S})) = \lim_{L\to\infty} \frac{H(L)}{L}$$

Directly from process's $\epsilon\mathrm{M}$:

$$h_\mu\left(\Pr(\overleftrightarrow{S})\right) = h_\mu(\boldsymbol{\mathcal{S}})$$

# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$:

Entropy Rate given $\epsilon\mathrm{M}$:

$$h_\mu(\boldsymbol{\mathcal{S}}) = - \sum_{\mathcal{S} \in \boldsymbol{\mathcal{S}}} \Pr(\mathcal{S}) \sum_{s \in \mathcal{A}, \mathcal{S}' \in \boldsymbol{\mathcal{S}}} T^{(s)}_{\mathcal{S}\mathcal{S}'} \log_2 T^{(s)}_{\mathcal{S}\mathcal{S}'}$$

where $\Pr(\mathcal{S})$ is casual-state asymptotic probability.

Possible only due to $\epsilon\mathrm{M}$ unifilarity!

1-1 mapping between measurement sequences & internal paths.

# Measures of Complexity ...

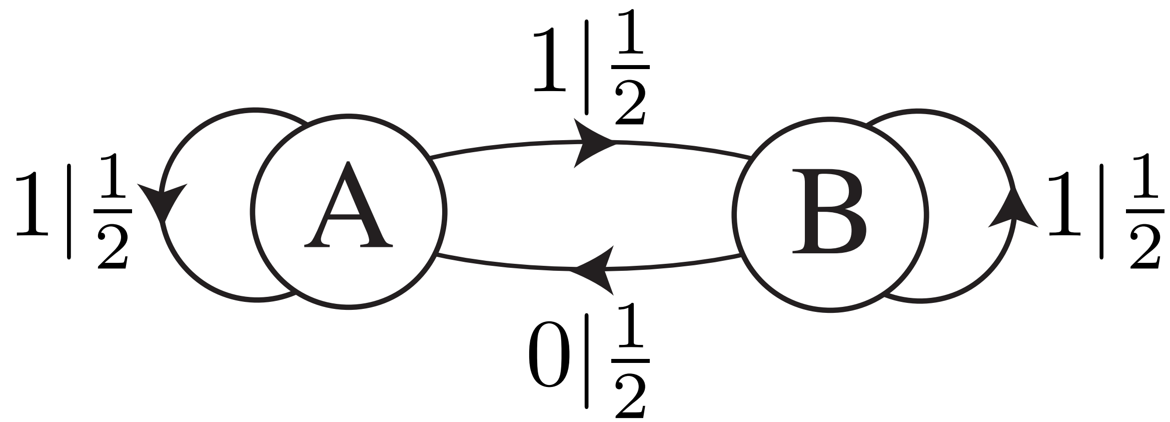
## Measures from the $\epsilon\mathrm{M}$:

### Entropy rate ...

Possible only due to unifilarity ...

What if you have a nonunifilar HMM for a process?

Consider SNS representation of a process:




$$\mathcal{B} = \{0, 1\}$$

$$T^{(0)} = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{pmatrix} \quad T^{(1)} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}$$

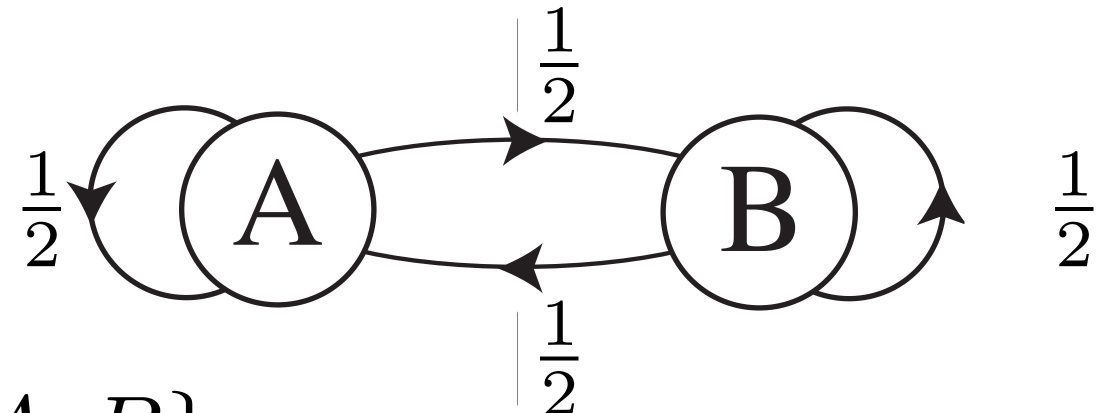
# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$:

Entropy rate ...

Possible only due to unifilarity ...

Internal Markov chain:




Internal (= Fair Coin): $\mathcal{A} = \{A, B\}$

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \pi_V = \left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

Entropy rate?

$$\begin{aligned} h_\mu(\mathrm{SNS}) &= -\sum_{v \in \mathcal{A}} \Pr(v) \sum_{v' \in \mathcal{A}} \Pr(v'|v) \log_2 \Pr(v'|v) \\ &= 1 \end{aligned}$$

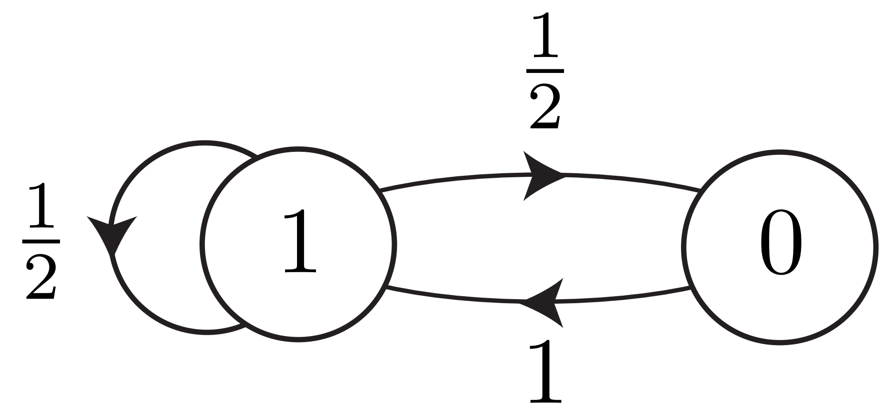
# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$:

### Entropy rate ...

### Possible only due to unifilarity ...

But support process is GMS




No consecutive 0s!

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \quad \pi_V = (2/3, 1/3)$$

A restriction that lowers entropy rate.

$$\begin{aligned} h_\mu(\mathrm{GMS}) &= -\frac{2}{3} \sum_{s \in \mathcal{B}} \Pr(s|1) \log_2 \Pr(s|1) \\ &= \frac{2}{3} \end{aligned}$$

# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$:

Entropy rate ...

Possible only due to unifilarity ...

$$h_\mu(\mathrm{SNS}) \gg h_\mu(\mathrm{GMS})$$

“SNS entropy rate” larger than support process: No!

The SNS representation overestimates entropy rate: internal process more random than observed process.

Lesson: Cannot use nonunifilar HMM representation of process to calculate entropy rate.

So how to compute SNS process’s entropy rate?

## Measures from the $\epsilon\mathrm{M}$:

### Entropy rate ...

Lesson: Need $\epsilon\mathrm{M}$ to calculate entropy rate.

SNS example: Nontrivial, countably infinite $\epsilon\mathrm{M}$.

$$h_\mu \approx 0.6778 \text{ bits/symbol}$$

Curious:

Even to estimate a process's intrinsic randomness,
need to infer its structure.

## Measures from the $\epsilon$M...

### Statistical Complexity of $\epsilon$M:

$$C_\mu(\boldsymbol{\mathcal{S}}) = - \sum_{\mathcal{S} \in \boldsymbol{\mathcal{S}}} \Pr(\mathcal{S}) \log_2 \Pr(\mathcal{S})$$

where $\Pr(\mathcal{S})$ is casual-state asymptotic probability.

### Meaning:

Shannon information in the causal states.

# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$...

Statistical Complexity of a Process:

$$C_\mu \left( \Pr(\overleftrightarrow{S}) \right) = C_\mu(\boldsymbol{\mathcal{S}})$$

Meaning:

The amount of historical information a process stores.

The amount of structure in a process.

## Measures from the $\epsilon\mathrm{M}$...

Excess Entropy: Three versions, all equivalent for 1D processes

$$\mathbf{E} = \lim_{L \to \infty} [H(L) - h_\mu L]$$

$$\mathbf{E} = \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu]$$

$$\mathbf{E} = I[\overleftarrow{S}; \overrightarrow{S}]$$

How to get, given $\epsilon\mathrm{M}$?

Special cases: When $\epsilon\mathrm{M}$ is IID, periodic, or spin chain.

General case: Need a new framework ... next lectures.

# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$...

### Excess Entropy ...

$\epsilon\mathrm{M}$ is IID:

$$\mathbf{E} = 0 \qquad C_\mu = 0$$

$\epsilon\mathrm{M}$ is Period $P$:

$$\mathbf{E} = \log_2 P = C_\mu$$

$\epsilon\mathrm{M}$ is range-R spin chain:

$$\mathbf{E} = H(R) - Rh_\mu$$

$$\mathbf{E} = C_\mu - Rh_\mu$$

# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$...

### Excess Entropy ...

Typically for Markov Chains:

$$\mathbf{E} < C_\mu$$

What can be said in general?

# Measures of Complexity ...

## Measures from the $\epsilon\mathrm{M}$ ...

### Bound on Excess Entropy:

$$\mathbf{E} \leq C_\mu$$

Proof sketch:

(1) $\mathbf{E} = I[\overrightarrow{S}; \overleftarrow{S}] = H[\overrightarrow{S}] - H[\overrightarrow{S} \mid \overleftarrow{S}]$

(2) Causal States: $H[\overrightarrow{S} \mid \overleftarrow{S}] = H[\overrightarrow{S} \,|\mathcal{S}]$

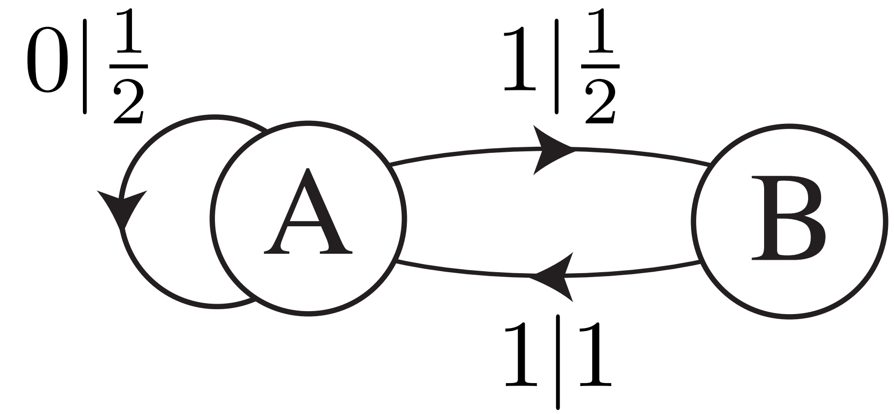
(3) $\mathbf{E} = H[\overrightarrow{S}] - H[\overrightarrow{S} \,|\mathcal{S}]$

$$= I[\overrightarrow{S}; \mathcal{S}]$$

$$= H[\mathcal{S}] - H[\mathcal{S}| \overrightarrow{S}]$$

$$\leq H[\mathcal{S}] = C_\mu$$

## Measures from the $\epsilon$M ...




Bound on Excess Entropy ...

But, the bound is saturated!

Even process:

$$T^{(0)} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$$

$$C_\mu = H(2/3) \approx 0.9182$$

$$\mathbf{E} \approx 0.9182$$

$$\pi_V = (2/3, 1/3)$$

When does this occur?

In general, need a new framework for answering this question.

Measures of Complexity ...

Measures from the $\epsilon\mathrm{M}$ ...

Bound on Excess Entropy ...

Consequence:

Can have $\mathbf{E} \to 0$ when $C_\mu \gg 1$. (Cryptographic limit)

Excess entropy is *not* the process's stored information.

$\mathbf{E}$ is the *apparent* information,
as revealed in *measurement sequences*.

Statistical complexity *is* stored information.

Measures from the $\epsilon\mathrm{M}$ ...

Bound on Excess Entropy ...

Executive Summary:

$C_\mu$ is the amount of information the process uses

to *communicate*

$\mathbf{E}$ bits of information from the past to the future.

## Measures from the $\epsilon\mathrm{M}$ ...

Bound on Excess Entropy: $\mathbf{E} \leq C_\mu$

Consequence:

The inequality is Why We Must Model.

Cannot simply use sequences as states.

There is internal structure not expressed by this.

# Information Diagrams for Processes

Process I-diagrams:

Process has an infinite number of RVs!

$$\Pr(\overleftrightarrow{X}) = \Pr(\ldots, X_{-2}, X_{-1}, X_0, X_1, X_2, \ldots)$$

Rather: $\Pr\left(\overleftrightarrow{X}\right) = \Pr\left(\overleftarrow{X}\,\overrightarrow{X}\right)$

Start with 2-variable I-diagram and whittle down:

Past as composite random variable: $\overleftarrow{X}$

Future as composite random variable: $\overrightarrow{X}$

Information measures:

$$H[\overleftarrow{X}] \quad H[\overrightarrow{X}] \quad H[\overrightarrow{X}, \overleftarrow{X}]$$

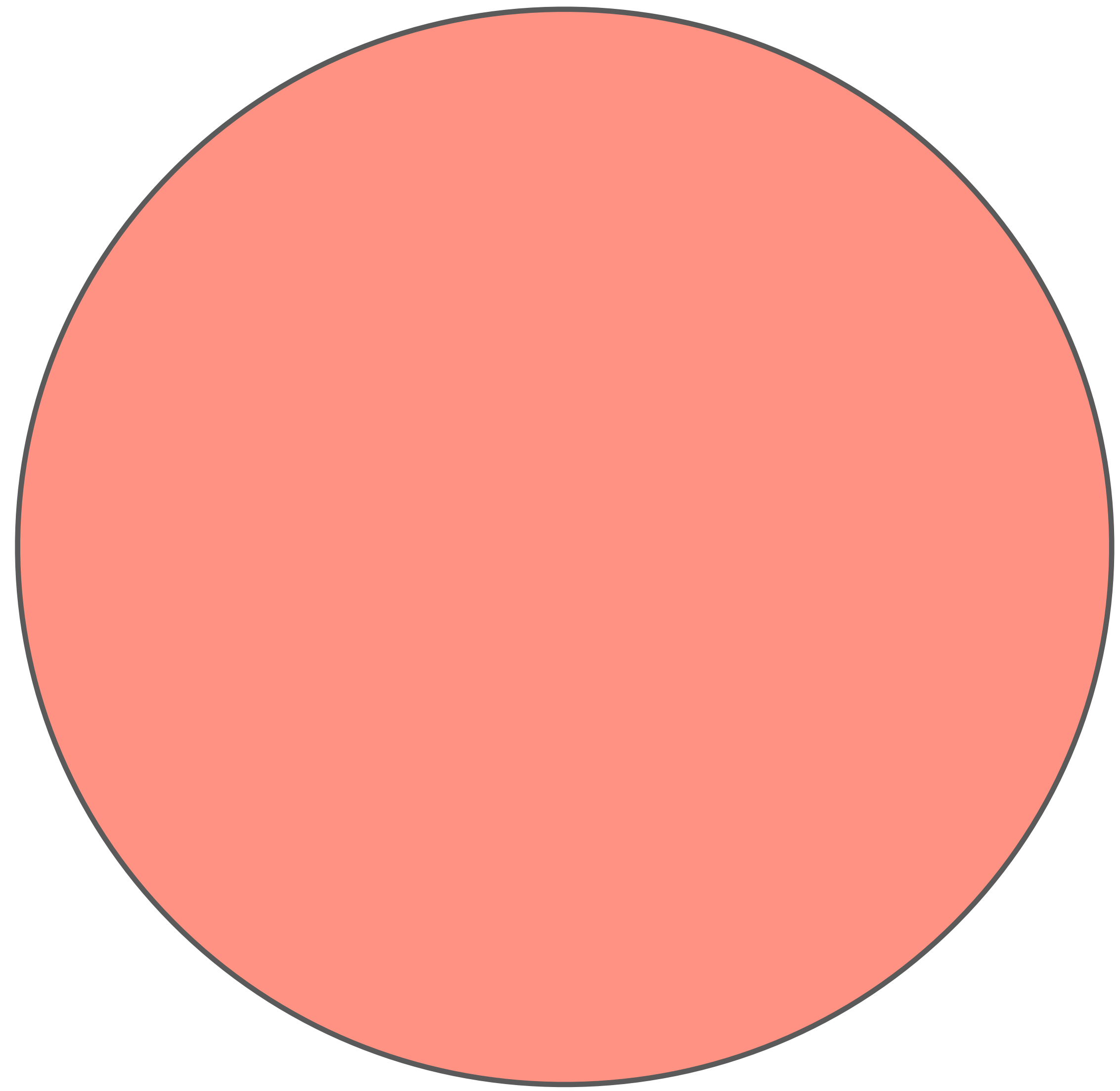
$$H[\overleftarrow{X}|\overrightarrow{X}] \quad H[\overrightarrow{X}|\overleftarrow{X}] \quad I[\overrightarrow{X}; \overleftarrow{X}] \quad H[\overrightarrow{X}|\overleftarrow{X}] + H[\overleftarrow{X}|\overrightarrow{X}]$$

There are $3 = 2^2 - 1$ atomic information measures:

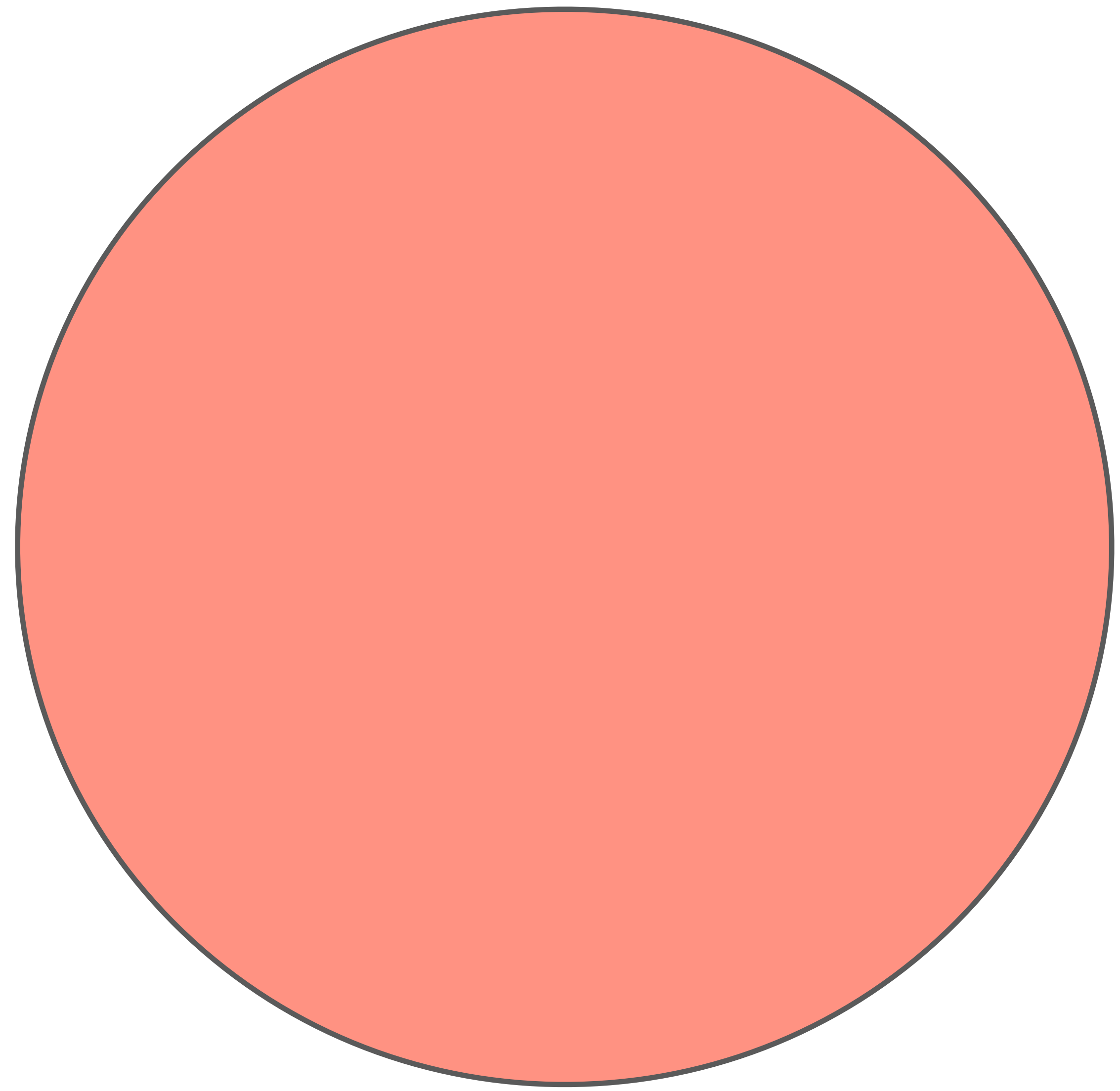
$$H[\overrightarrow{X}|\overleftarrow{X}] \quad H[\overleftarrow{X}|\overrightarrow{X}] \quad I[\overrightarrow{X}; \overleftarrow{X}]$$

# Information Diagrams for Processes

# Information Diagrams for Processes

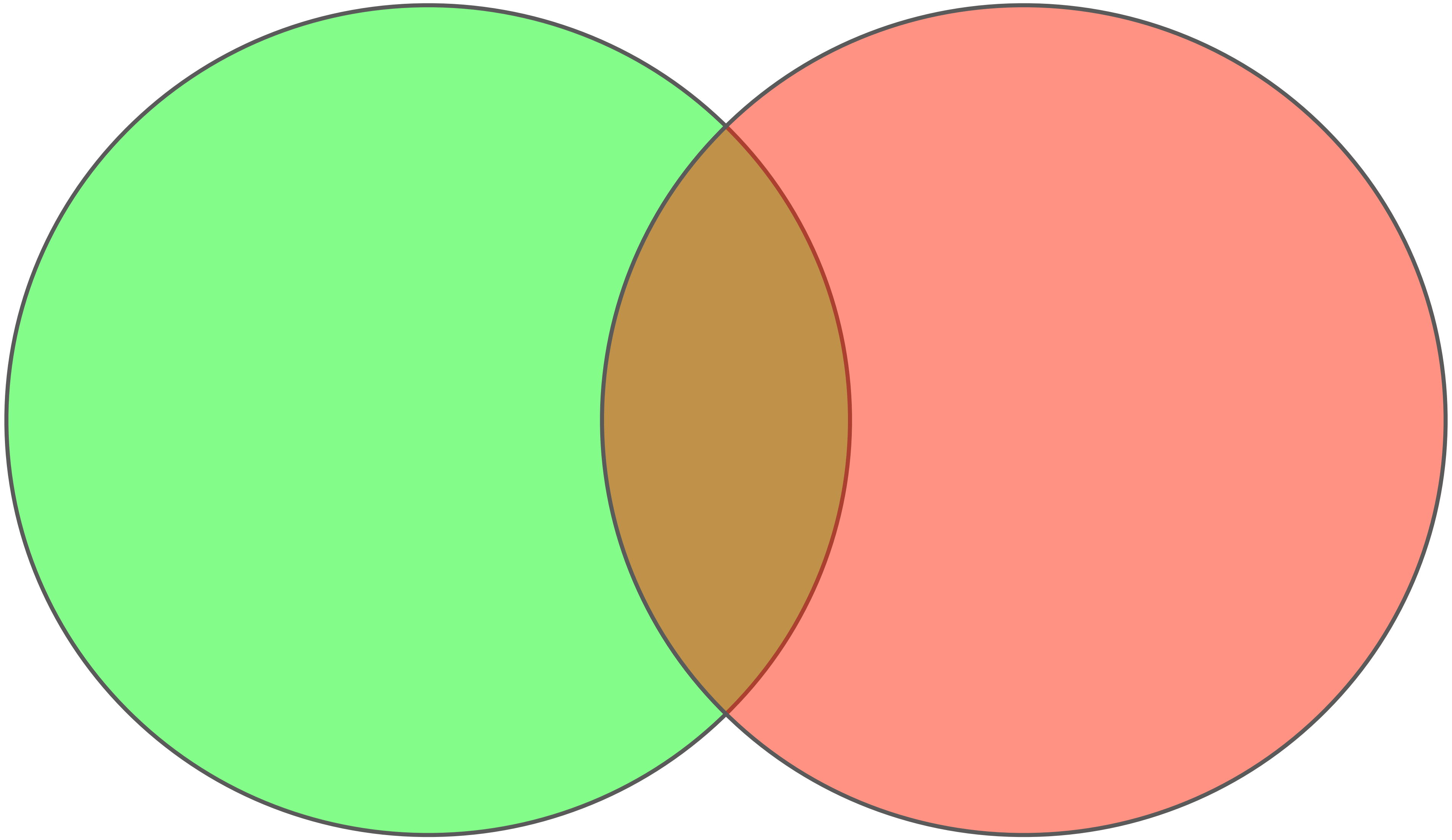


# Information Diagrams for Processes



$H[\overrightarrow{X}]$

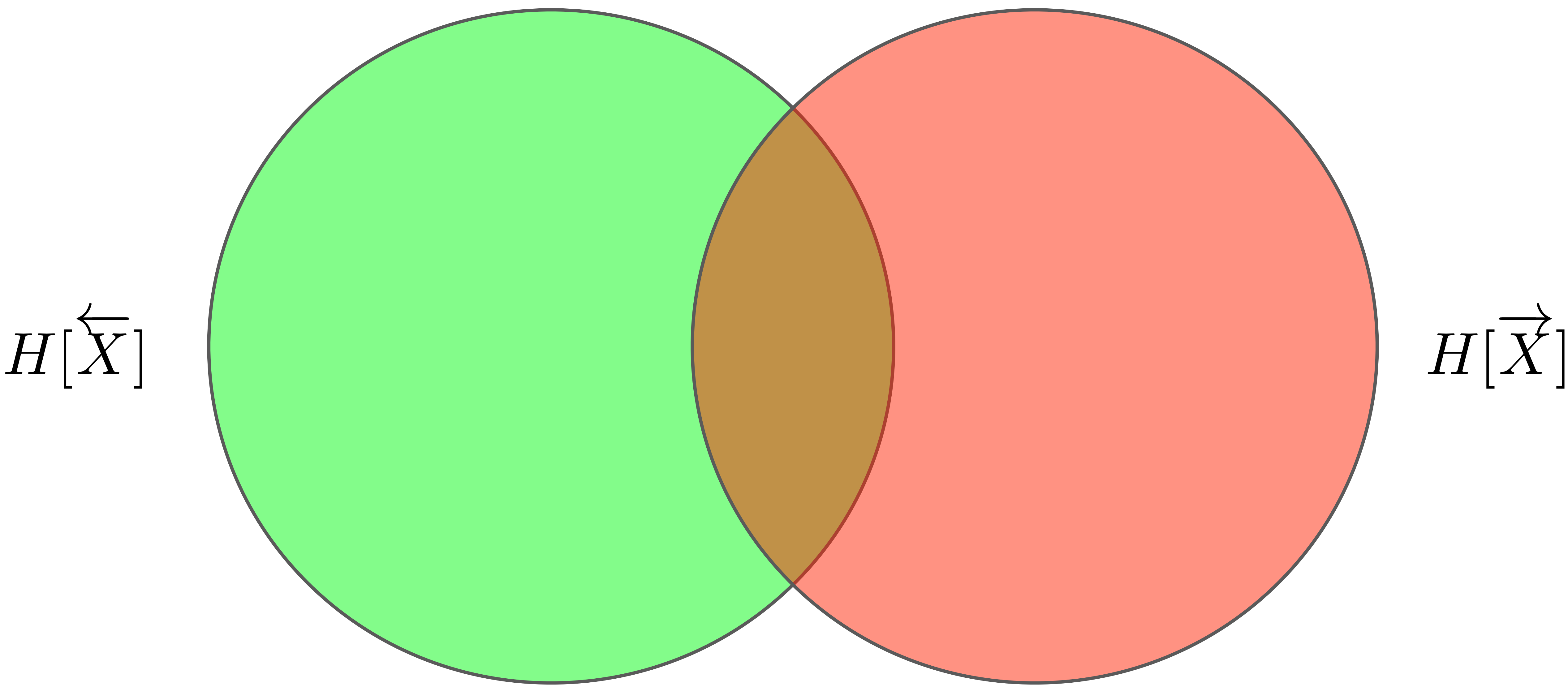
# Information Diagrams for Processes



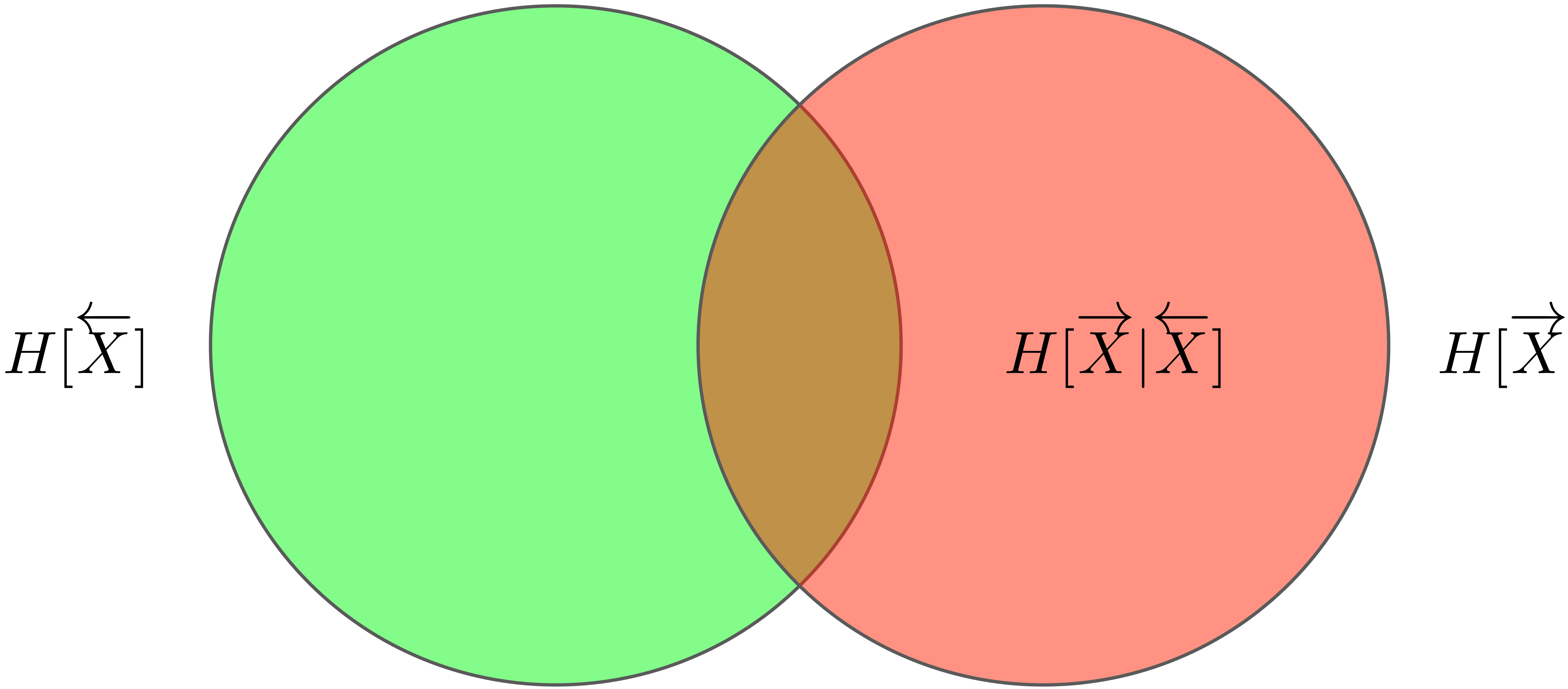
$H[\overrightarrow{X}]$

Complexity Lecture 2: Computational Mechanics (CSSS 2011); Jim Crutchfield

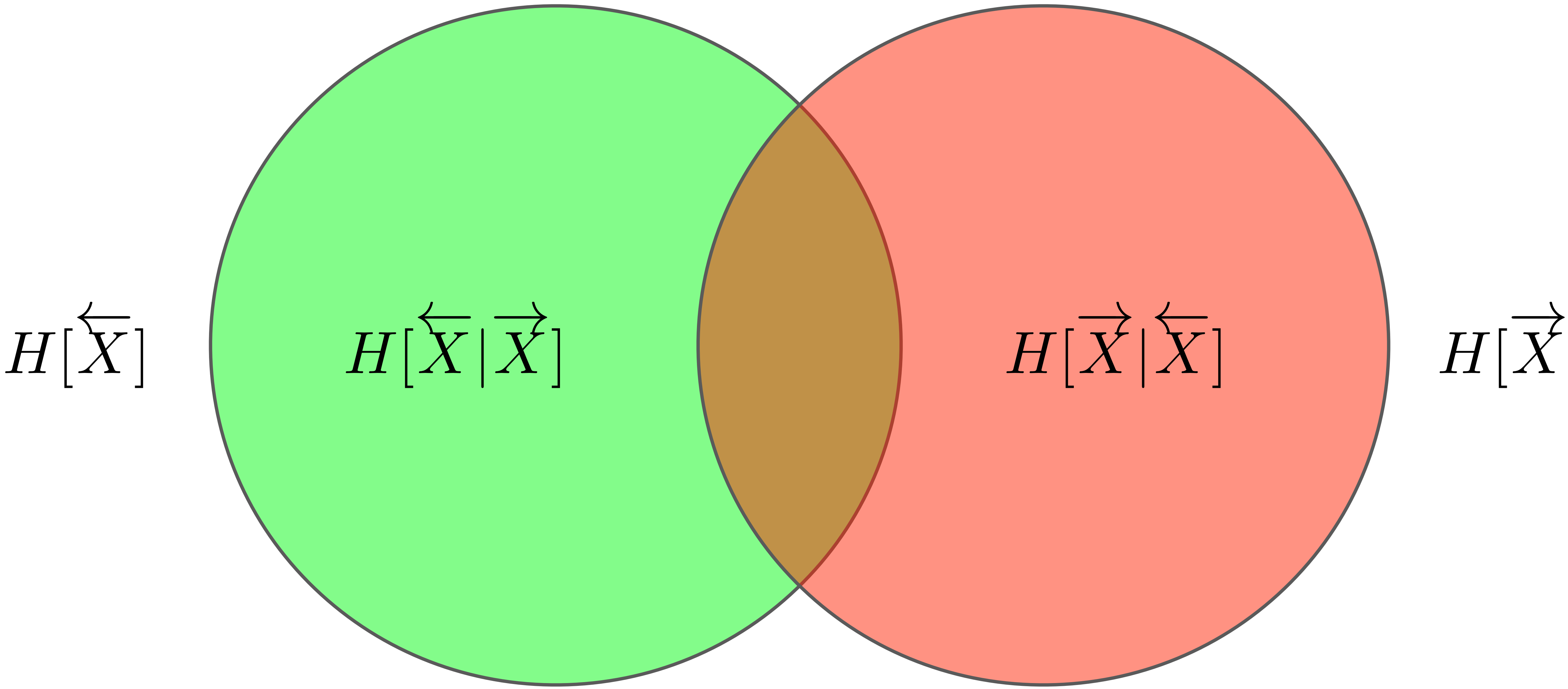
# Information Diagrams for Processes



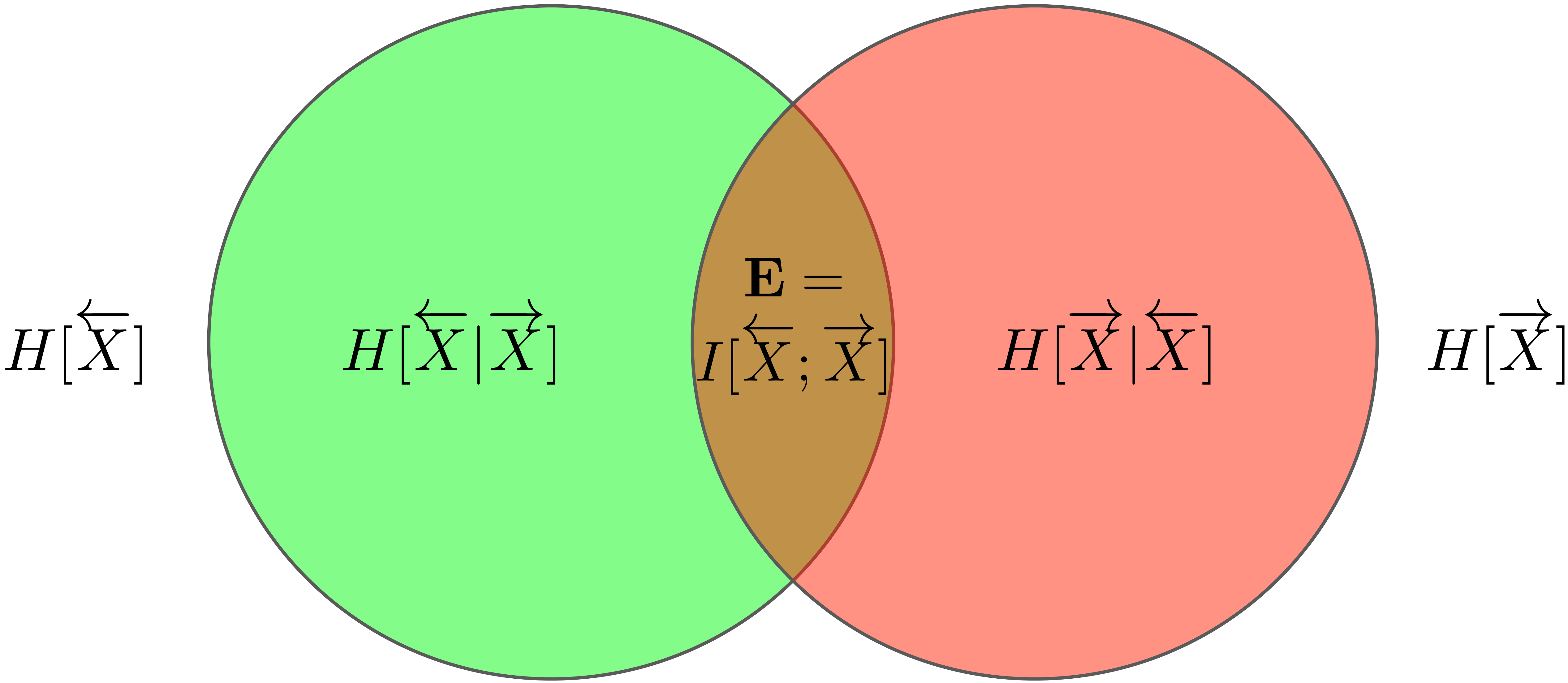
# Information Diagrams for Processes



# Information Diagrams for Processes



# Information Diagrams for Processes



# Information Diagrams for Processes

Process I-diagram using ε-machine:

Start with 3-variable I-diagram and whittle down:

- Past as composite random variable: $\overleftarrow{X}$
- Future as composite random variable: $\overrightarrow{X}$
- Causal states: $\mathcal{S} \in \boldsymbol{\mathcal{S}}$

Information measures:

$$H[\overleftarrow{X}] \quad H[\overrightarrow{X}] \quad H[\mathcal{S}] \quad \cdots \quad I[\overrightarrow{X};\overleftarrow{X};\mathcal{S}] \quad \cdots \quad H[\overrightarrow{X},\overleftarrow{X},\mathcal{S}]$$

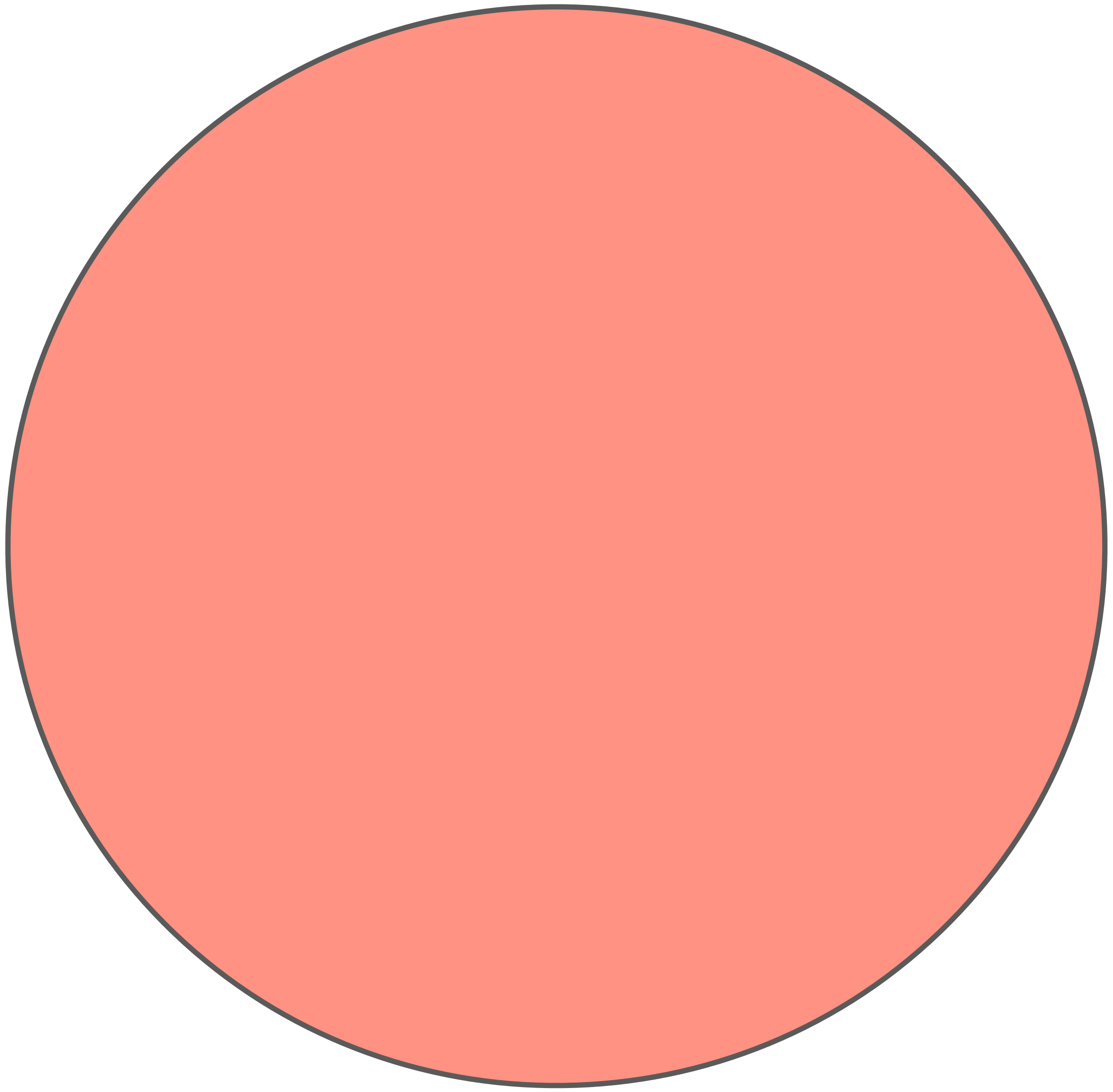
There are $8 = 2^3$ atomic information measures.

# Information Diagrams for Processes

## Process I-diagram using ε-machine ...

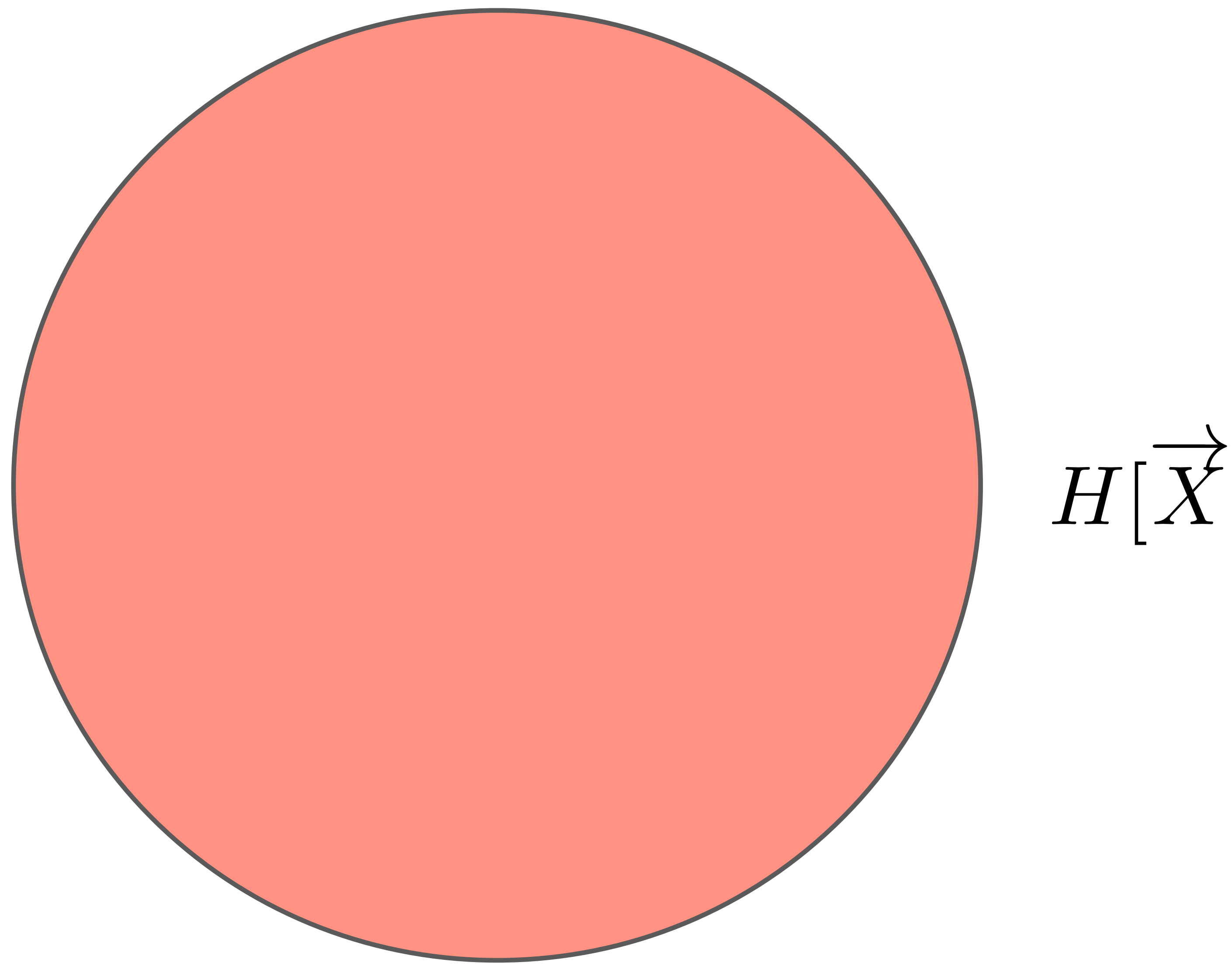
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



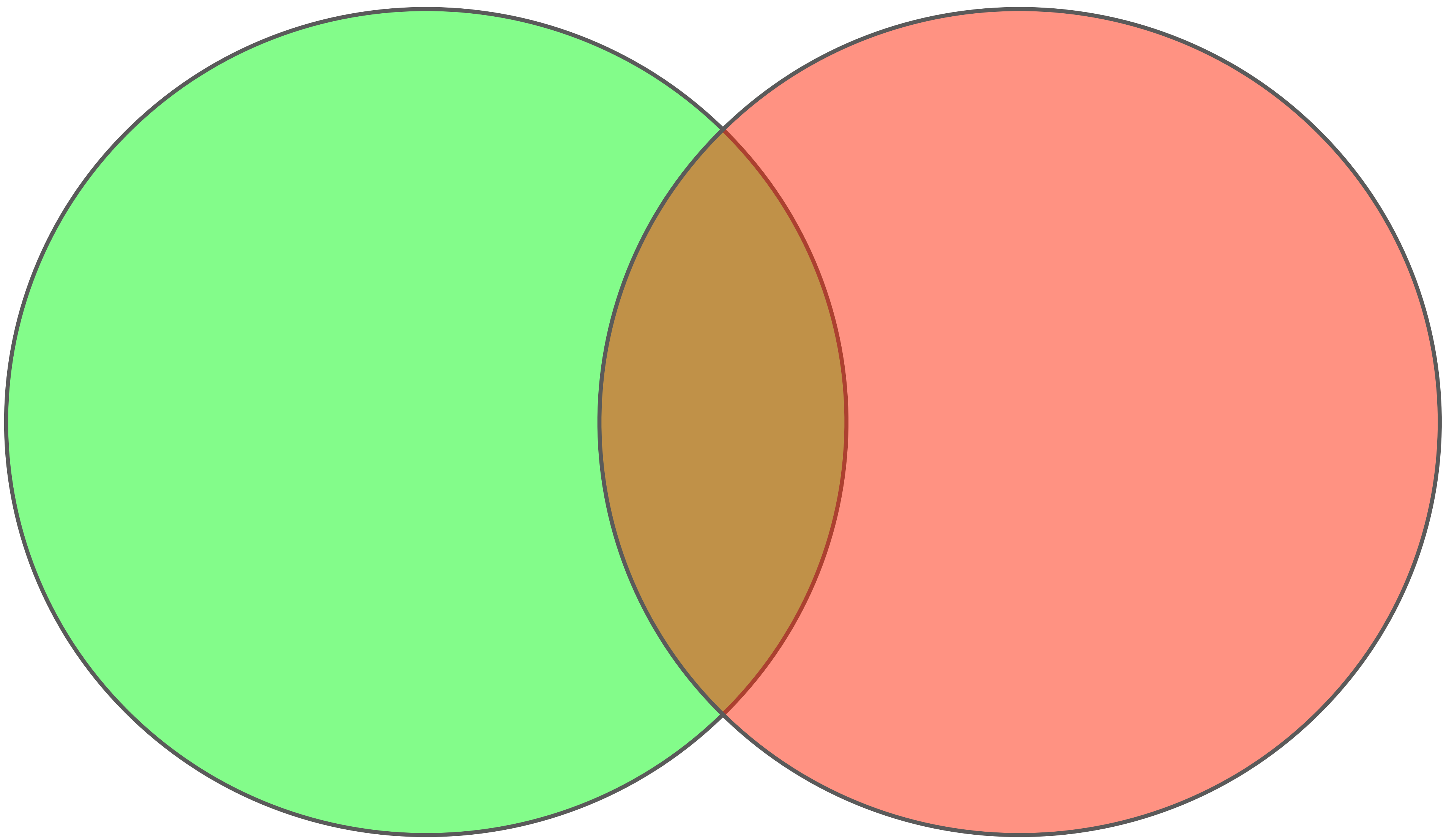
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



# Information Diagrams for Processes

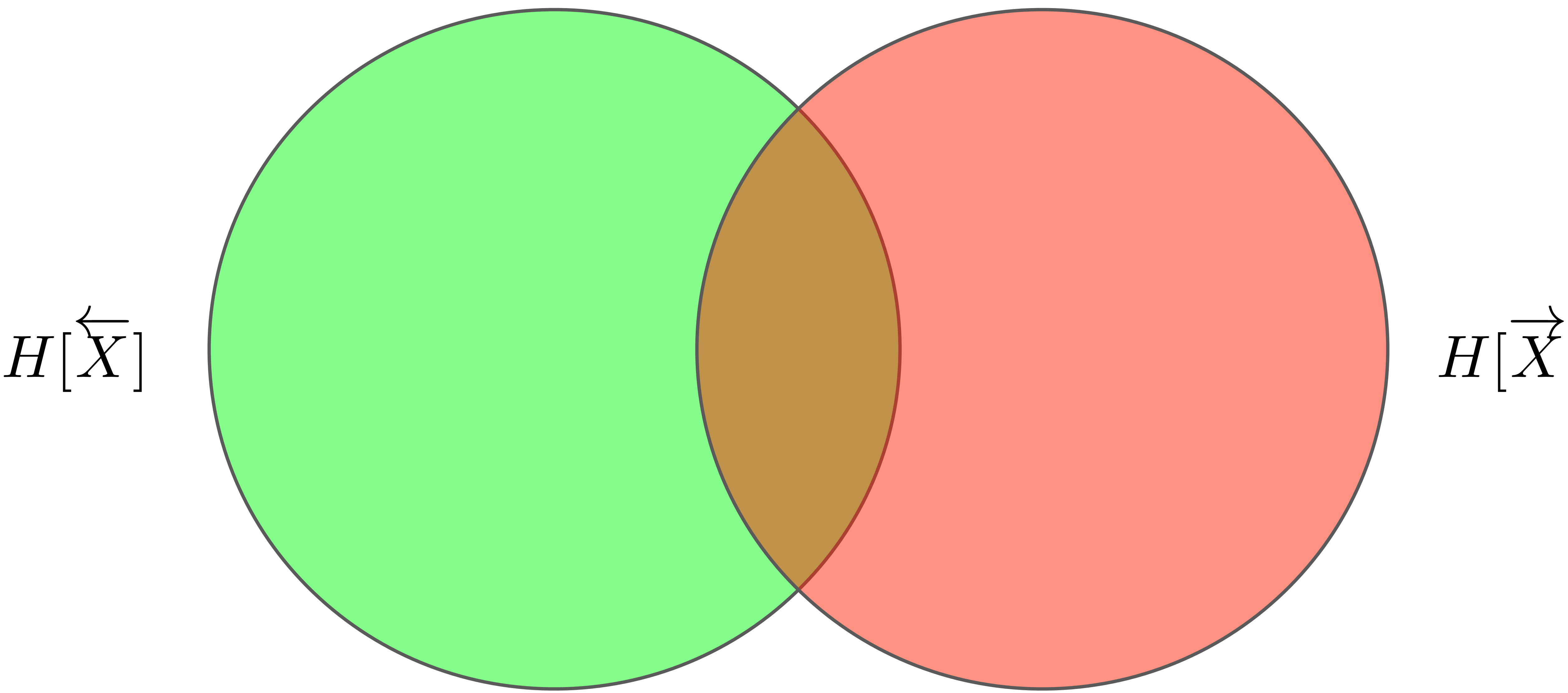
## Process I-diagram using ε-machine ...



$H[\overrightarrow{X}]$

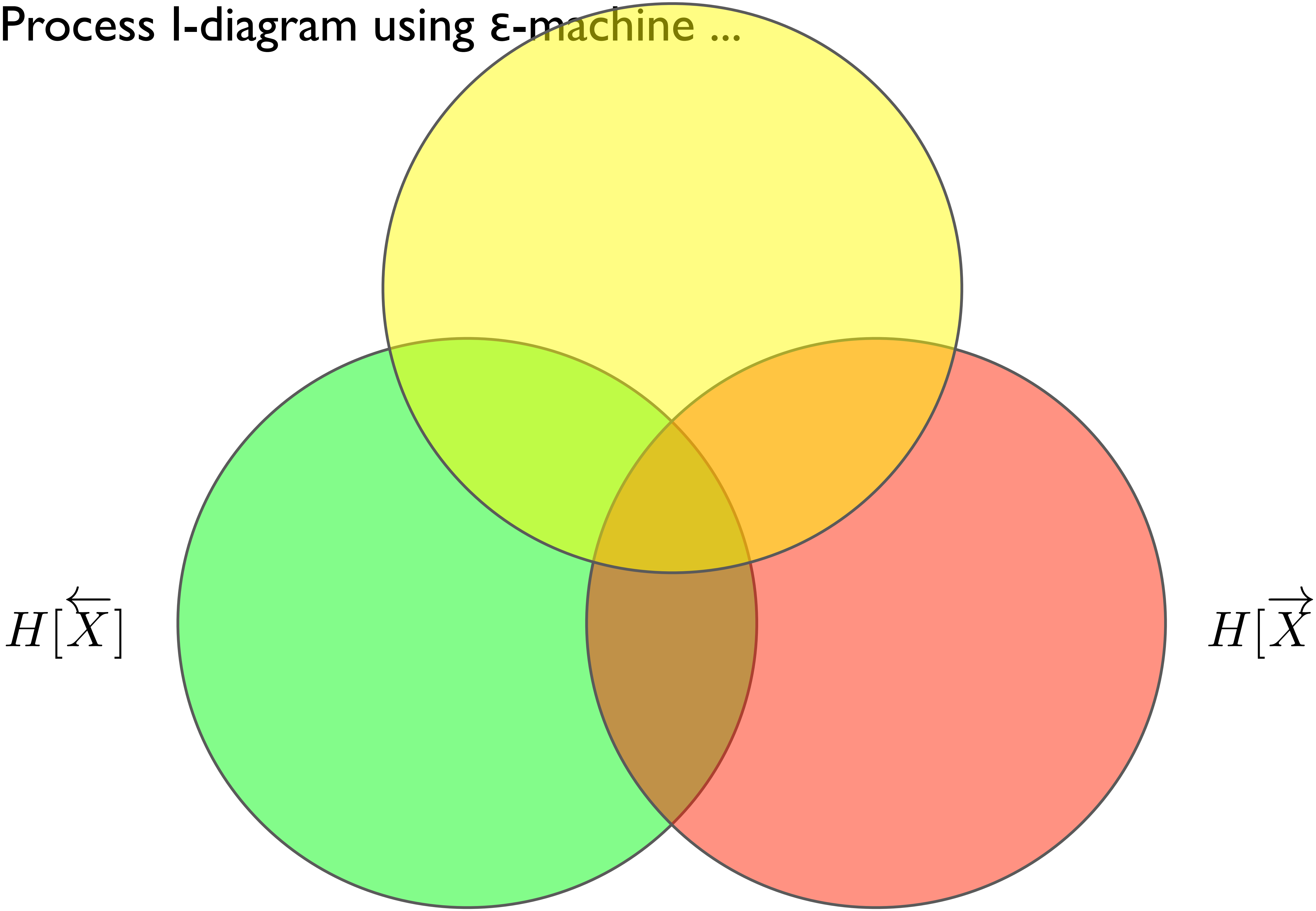
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



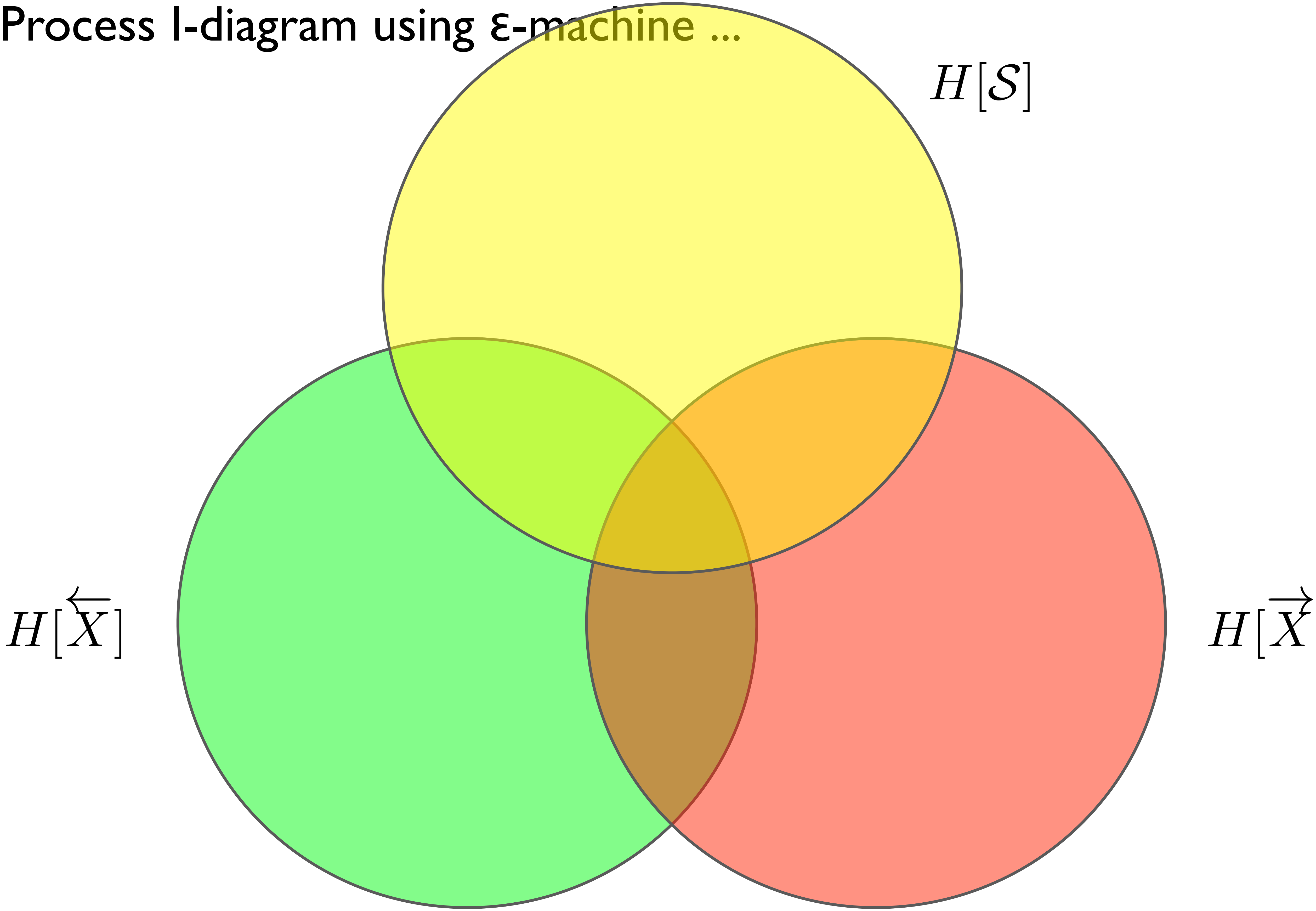
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



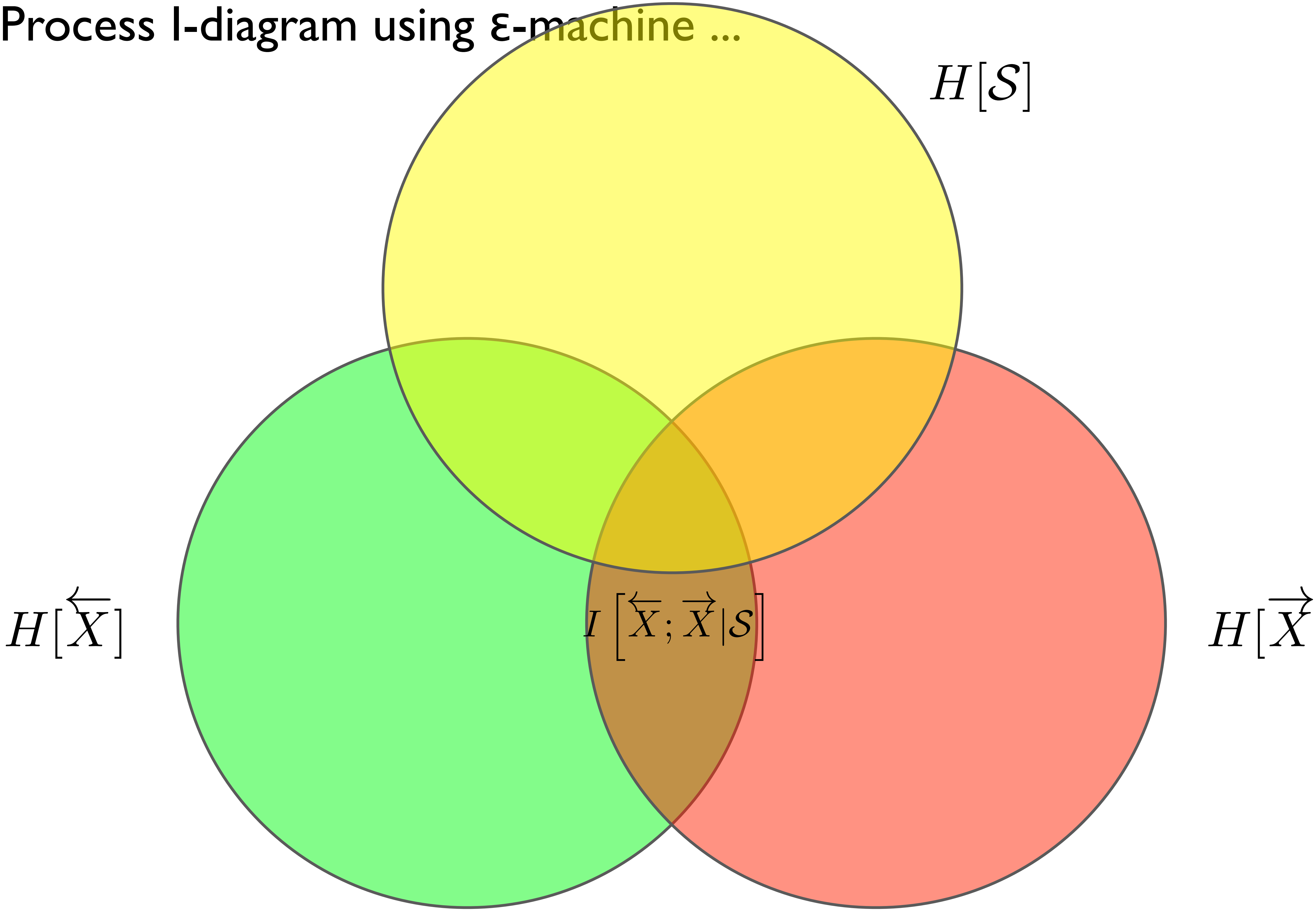
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



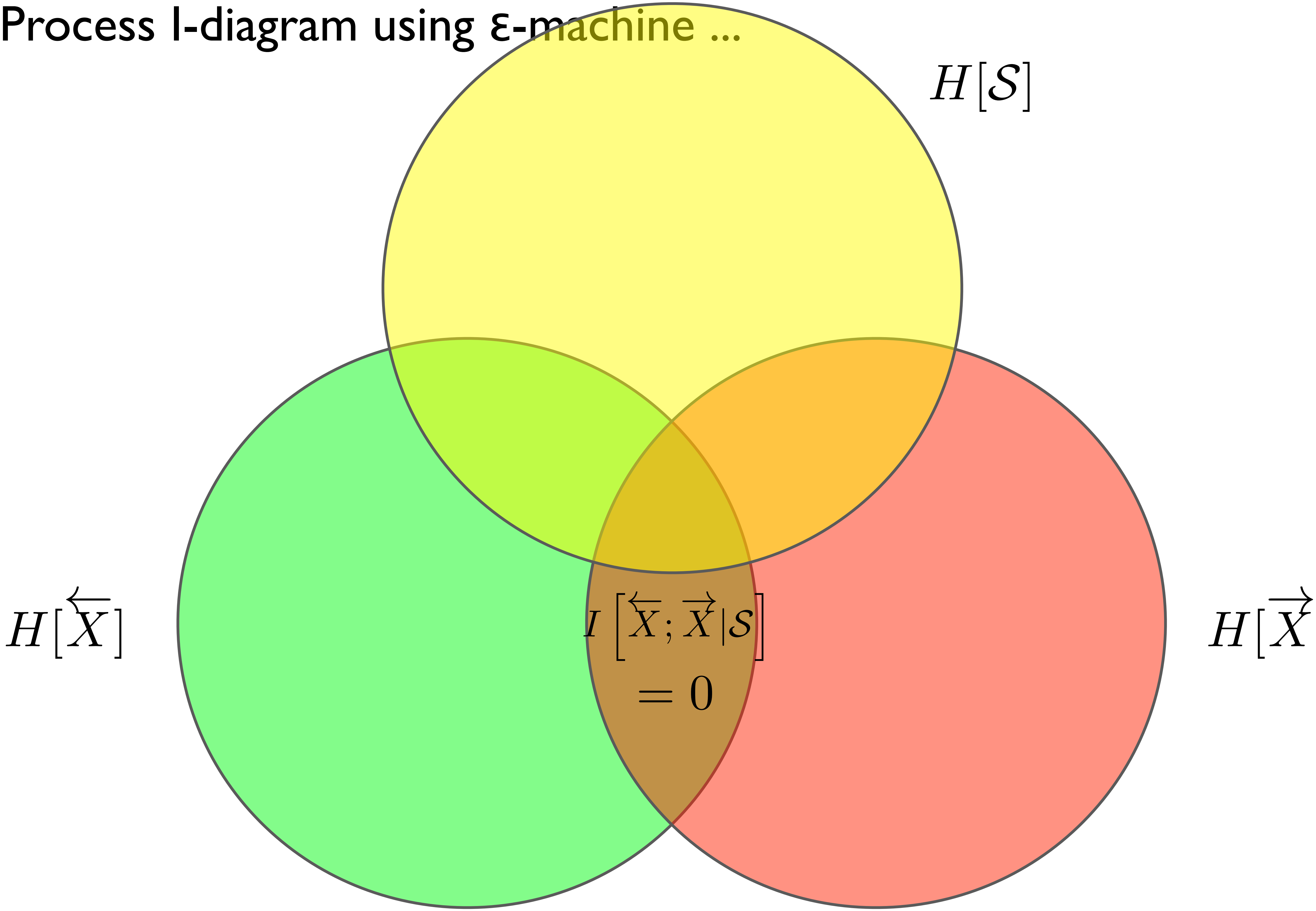
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



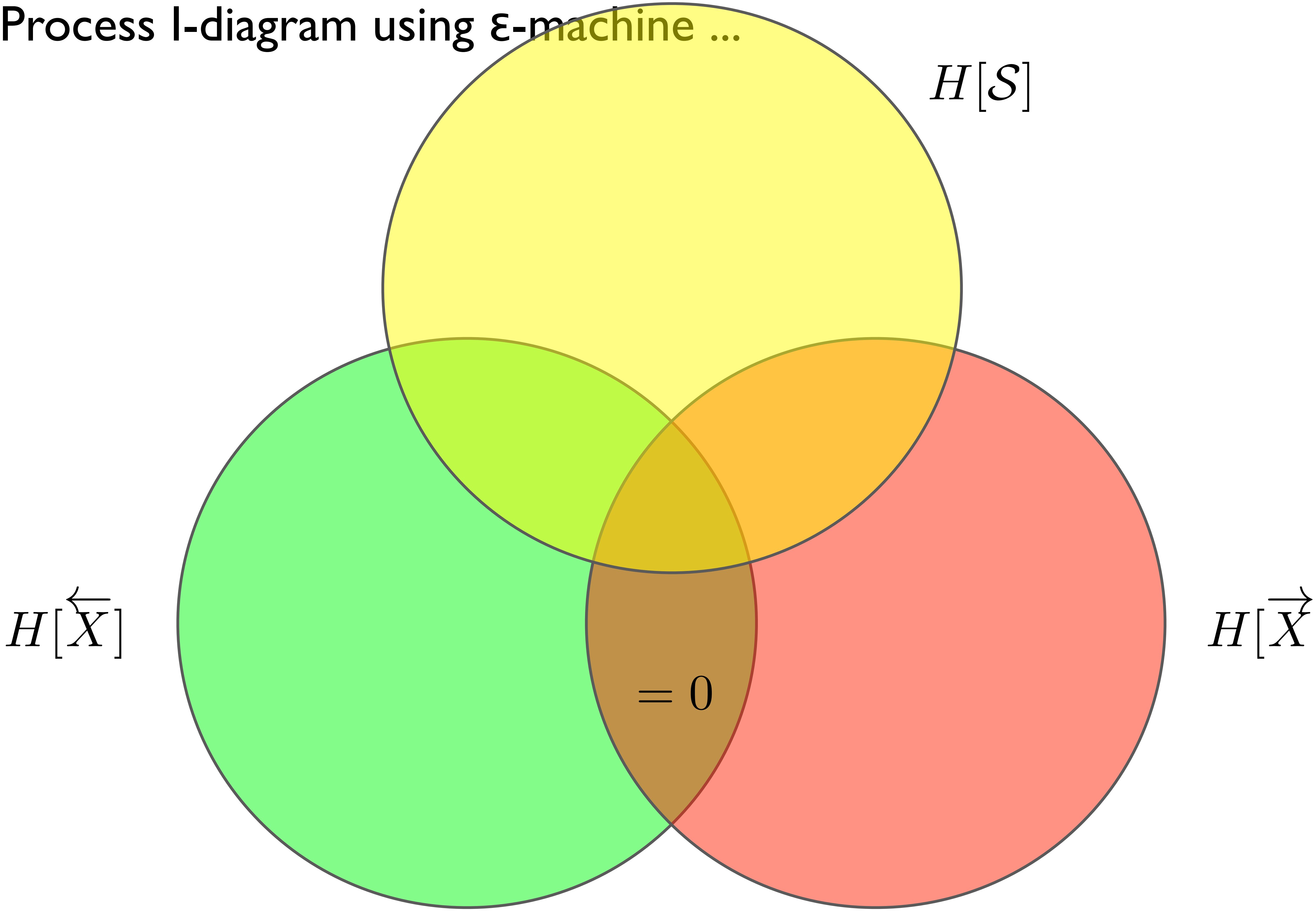
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



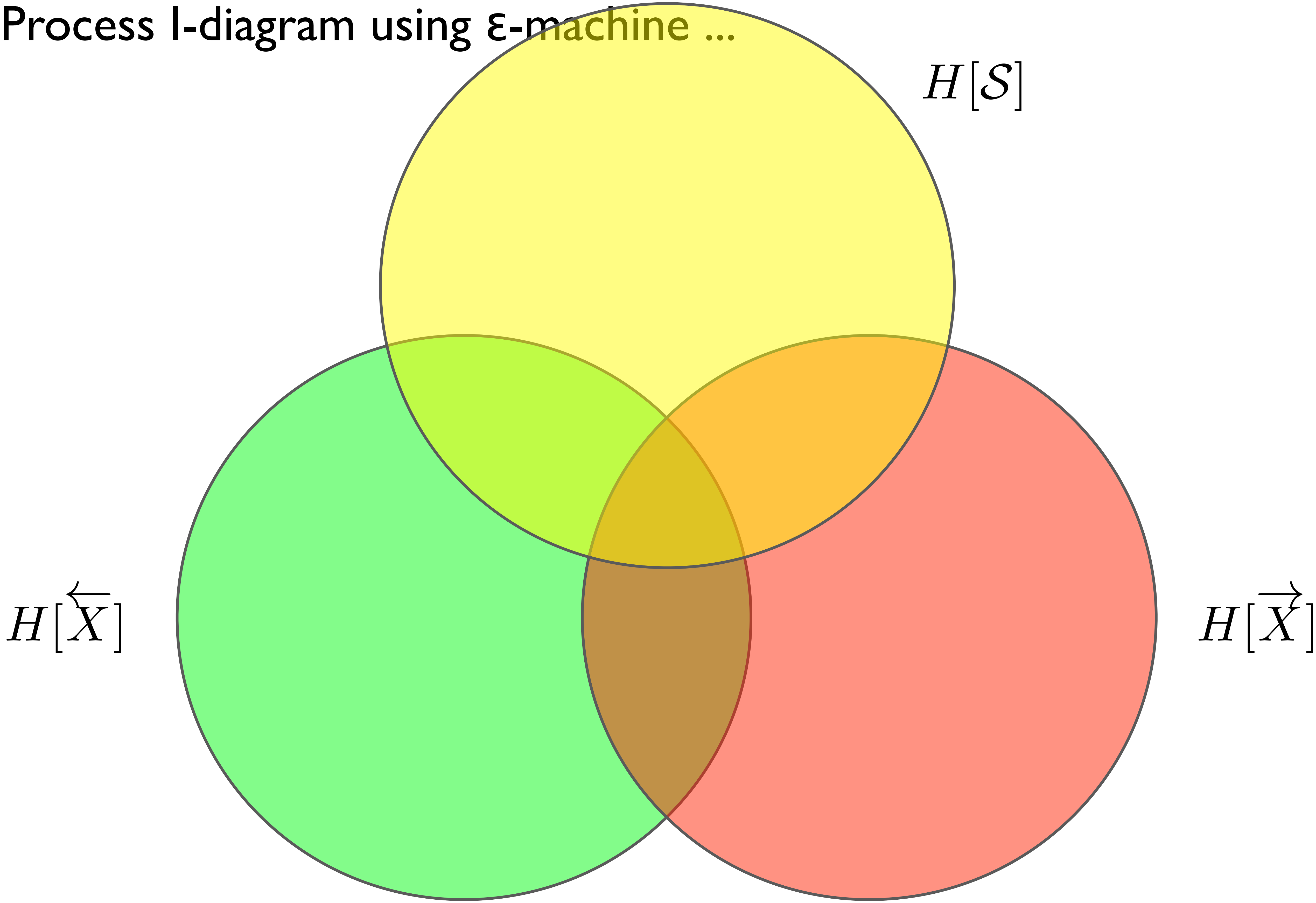
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



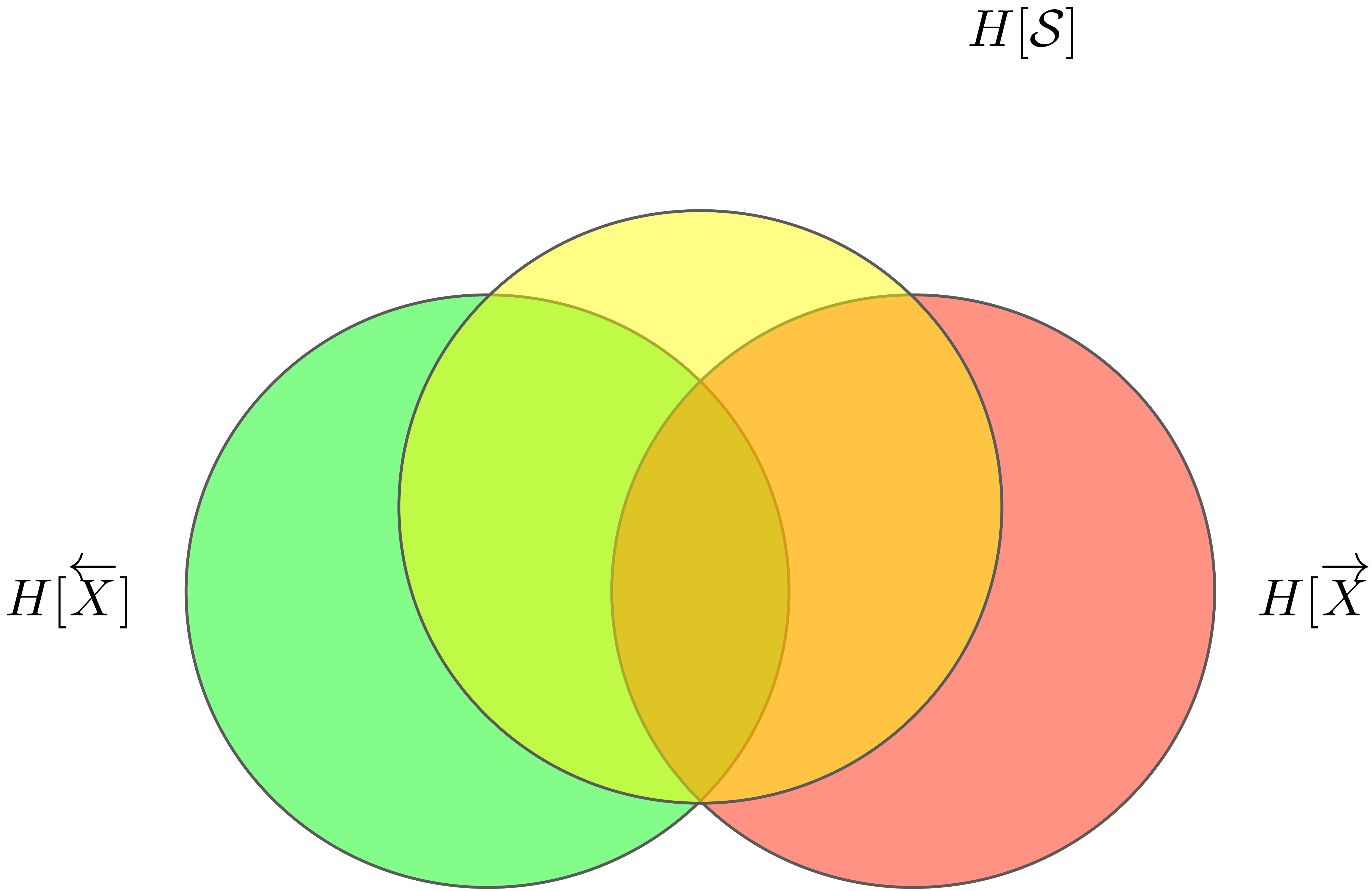
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



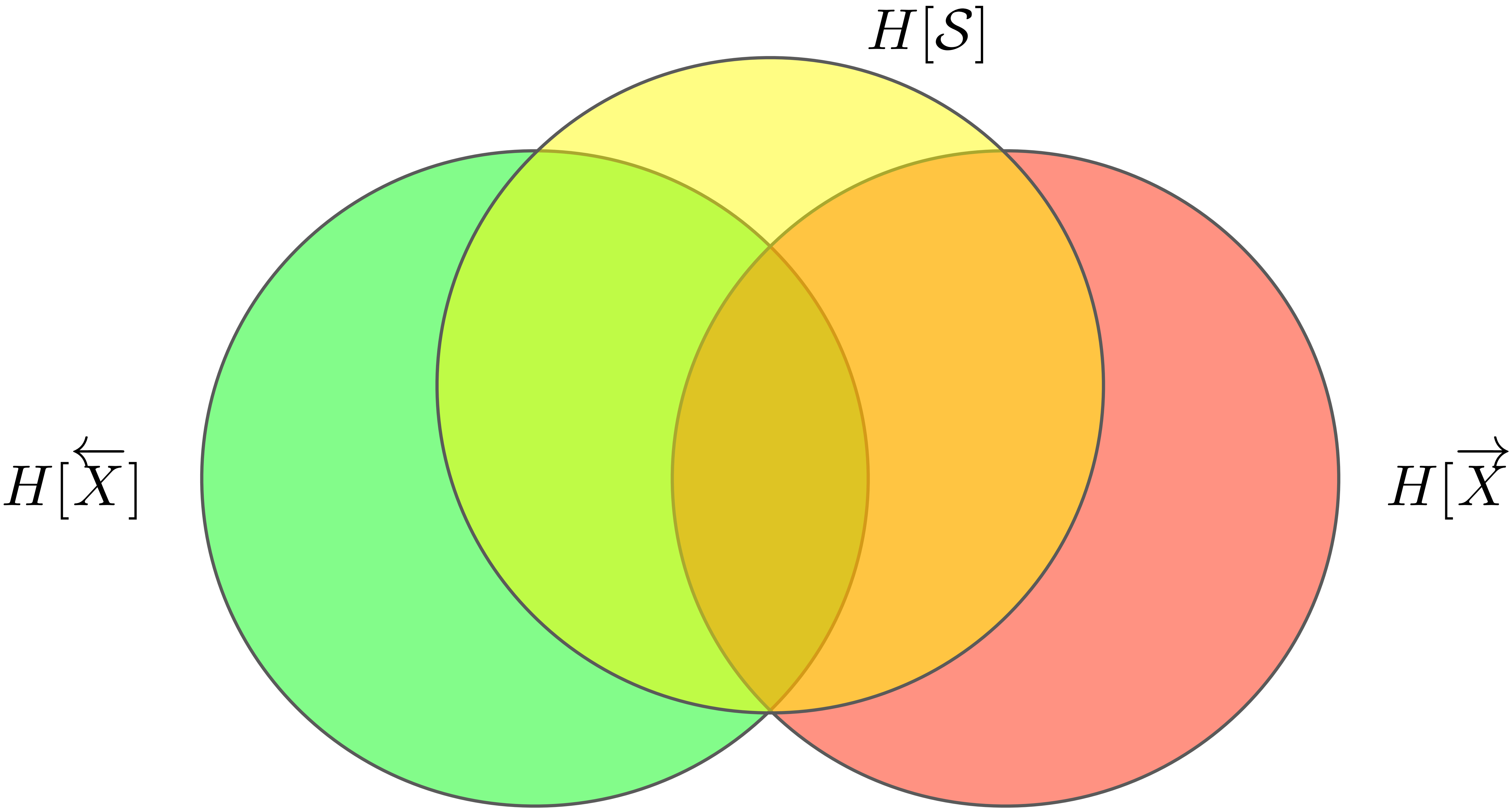
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



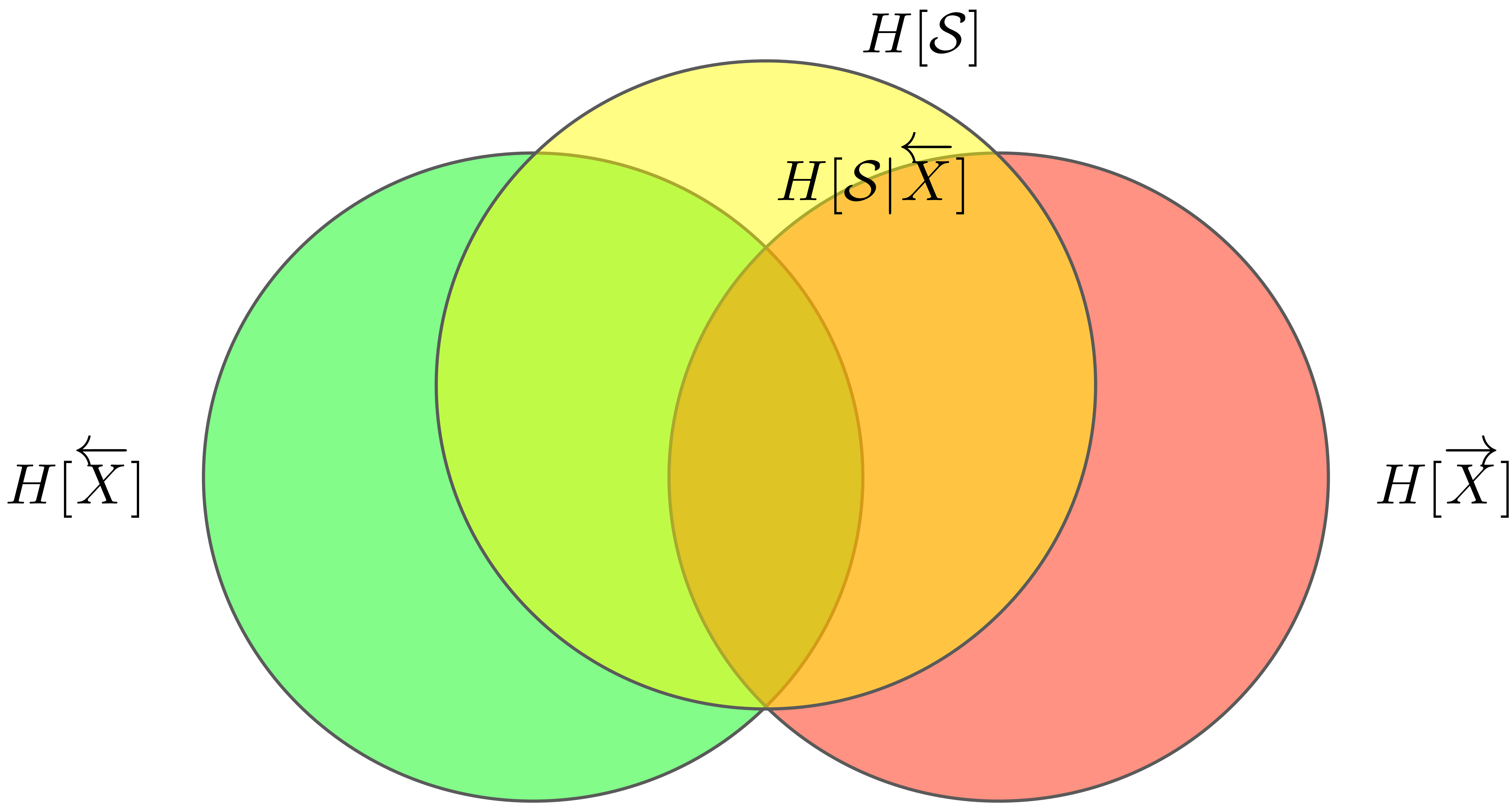
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



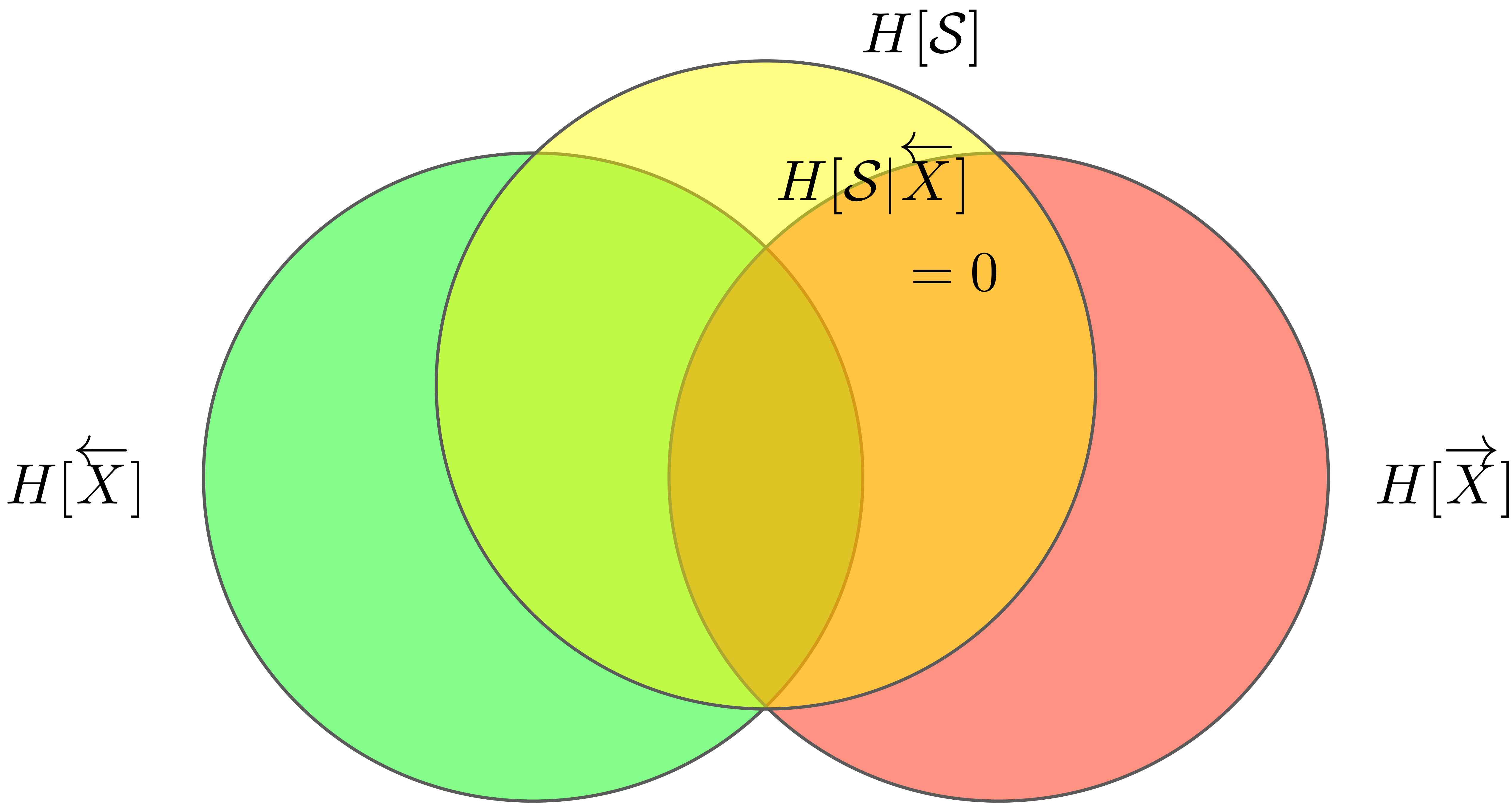
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



# Information Diagrams for Processes

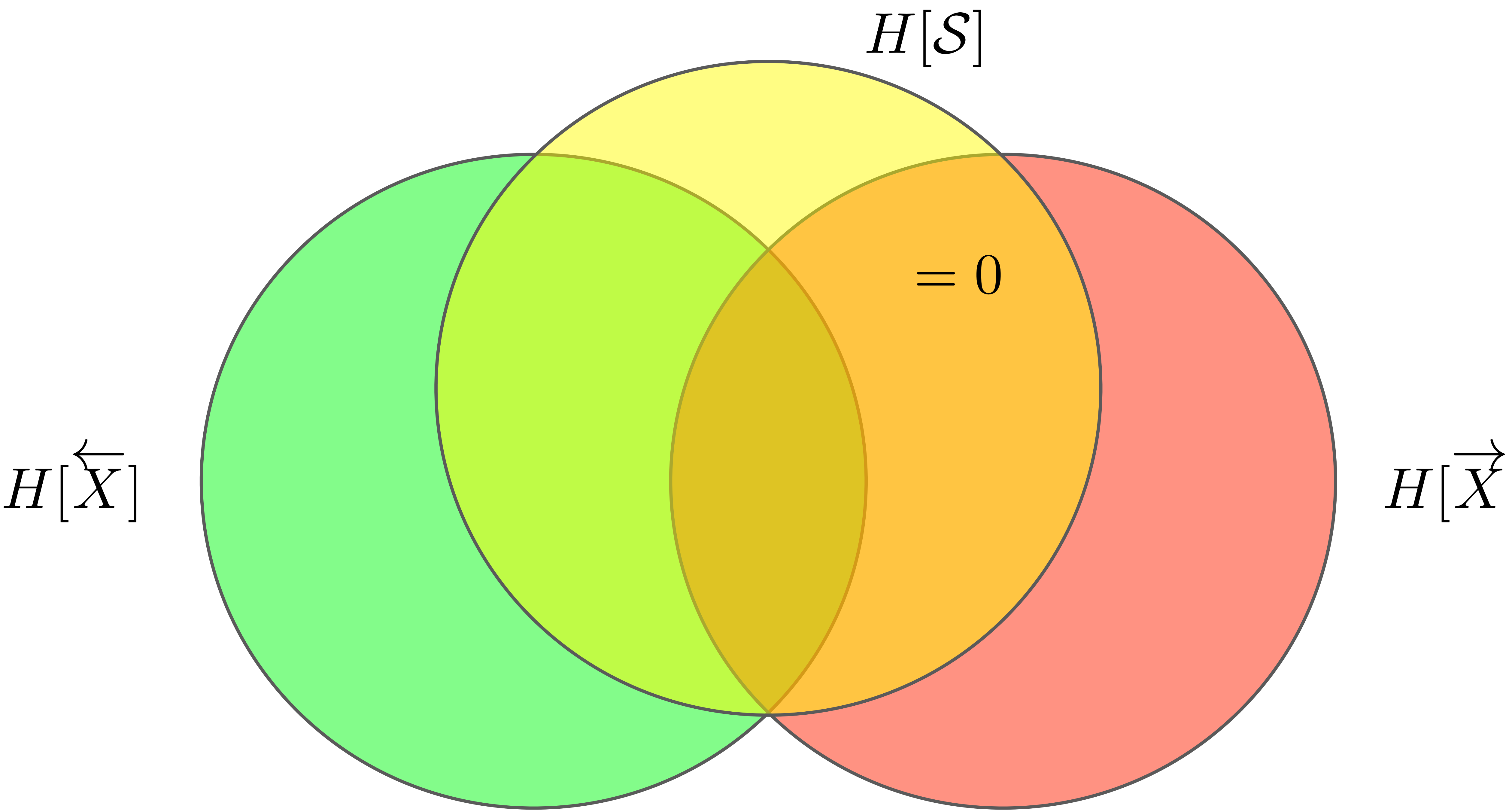
## Process I-diagram using ε-machine ...




Complexity Lecture 2: Computational Mechanics (CSSS 2011); Jim Crutchfield

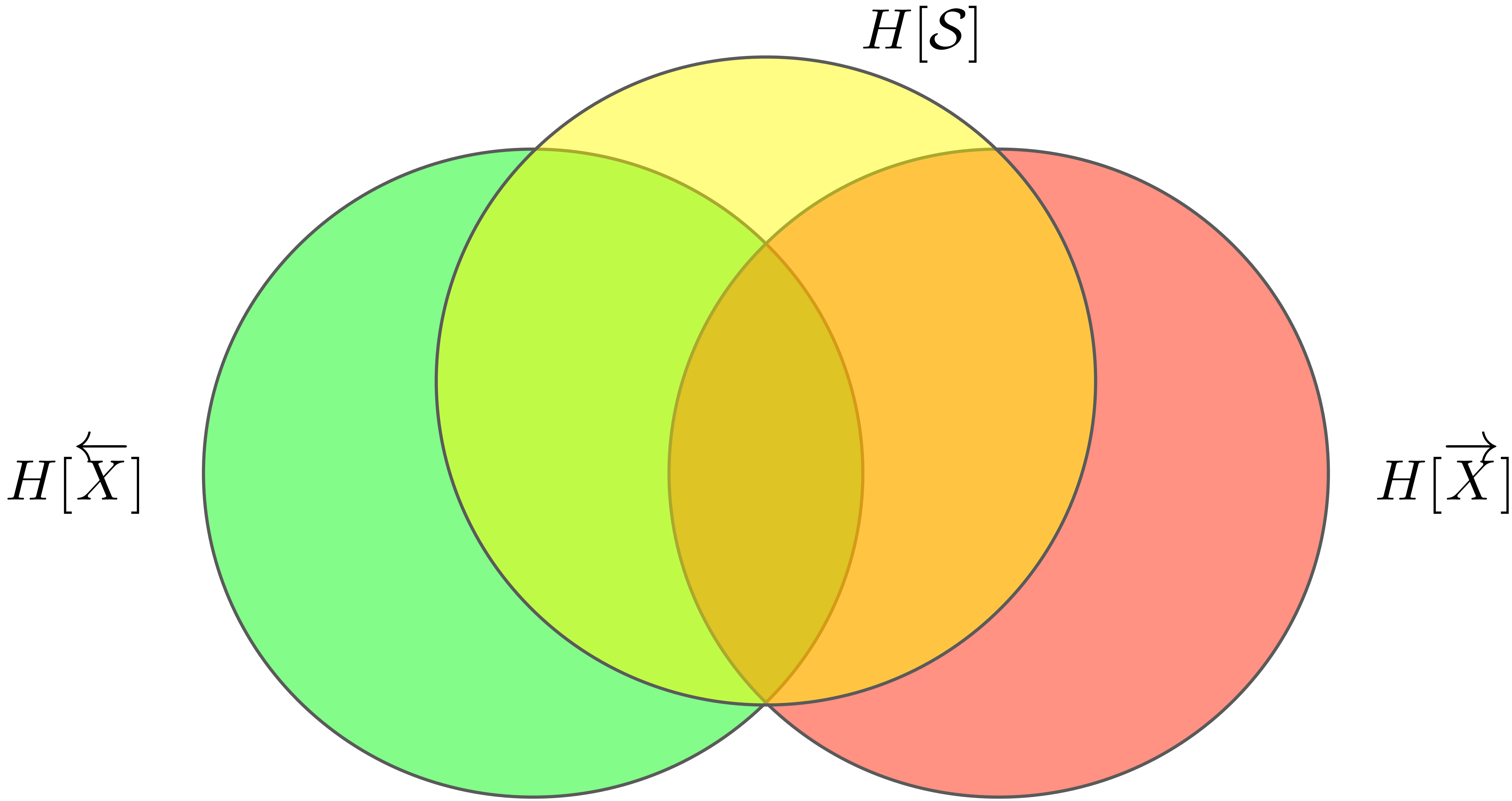
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



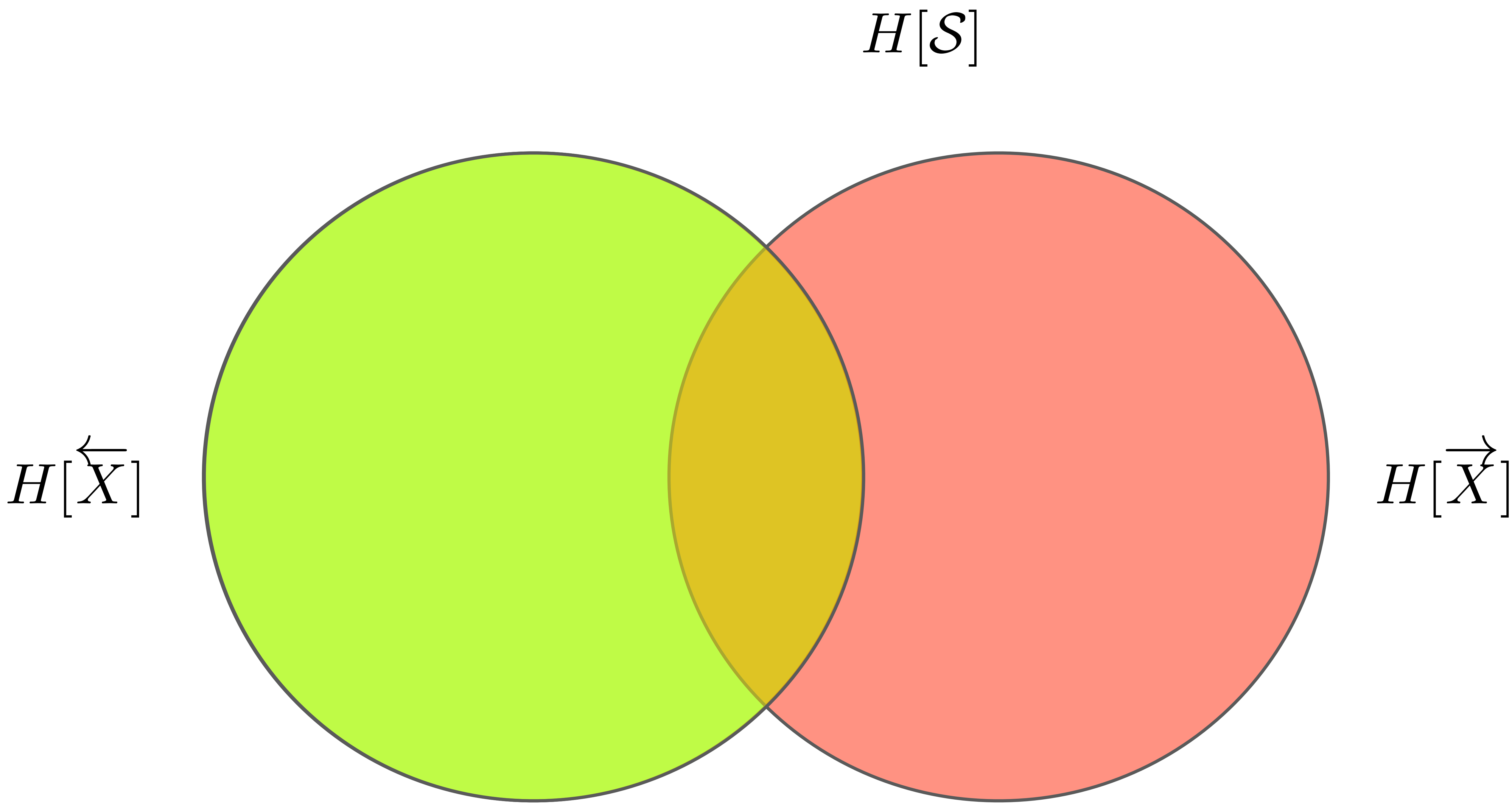
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



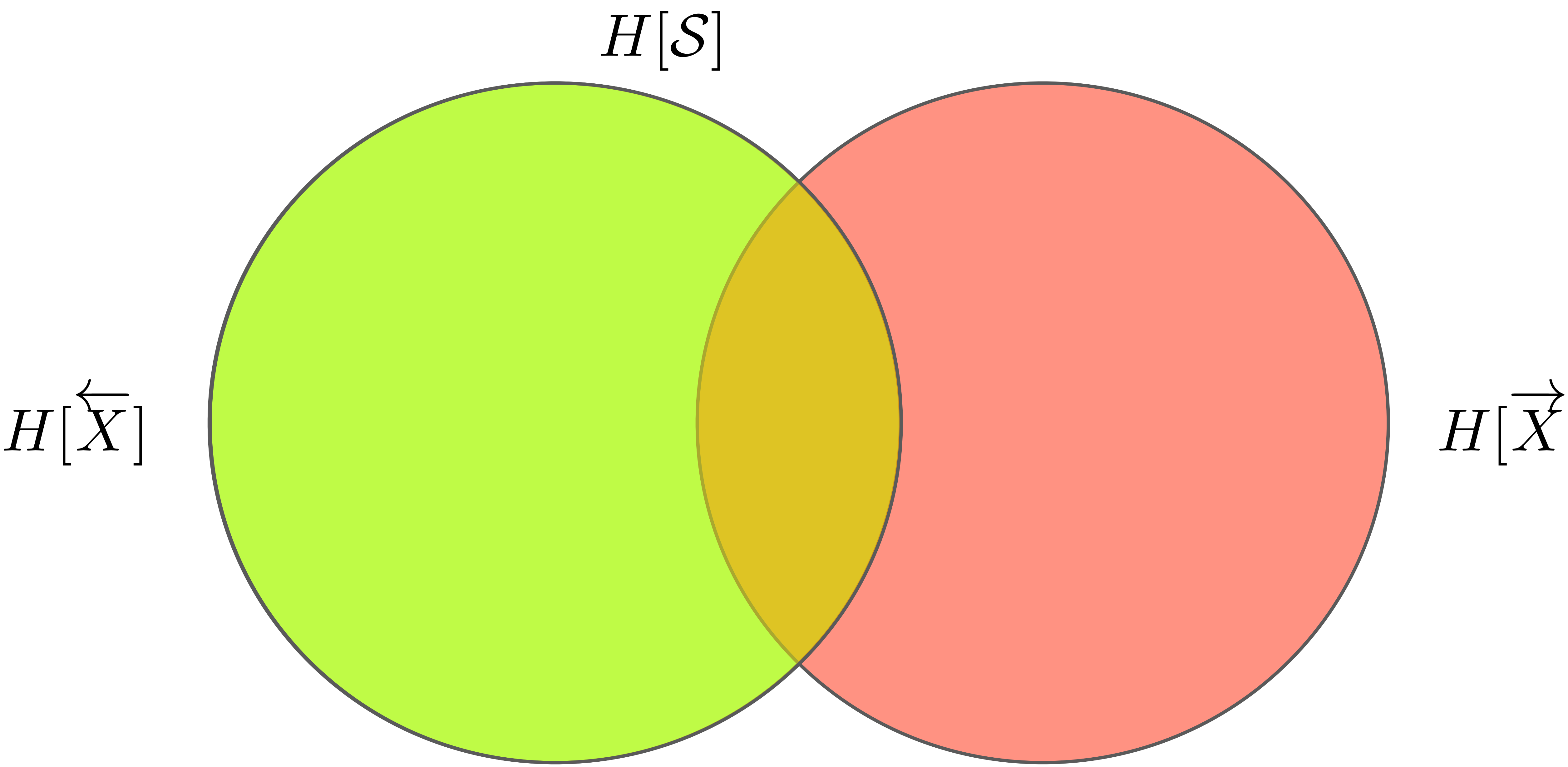
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



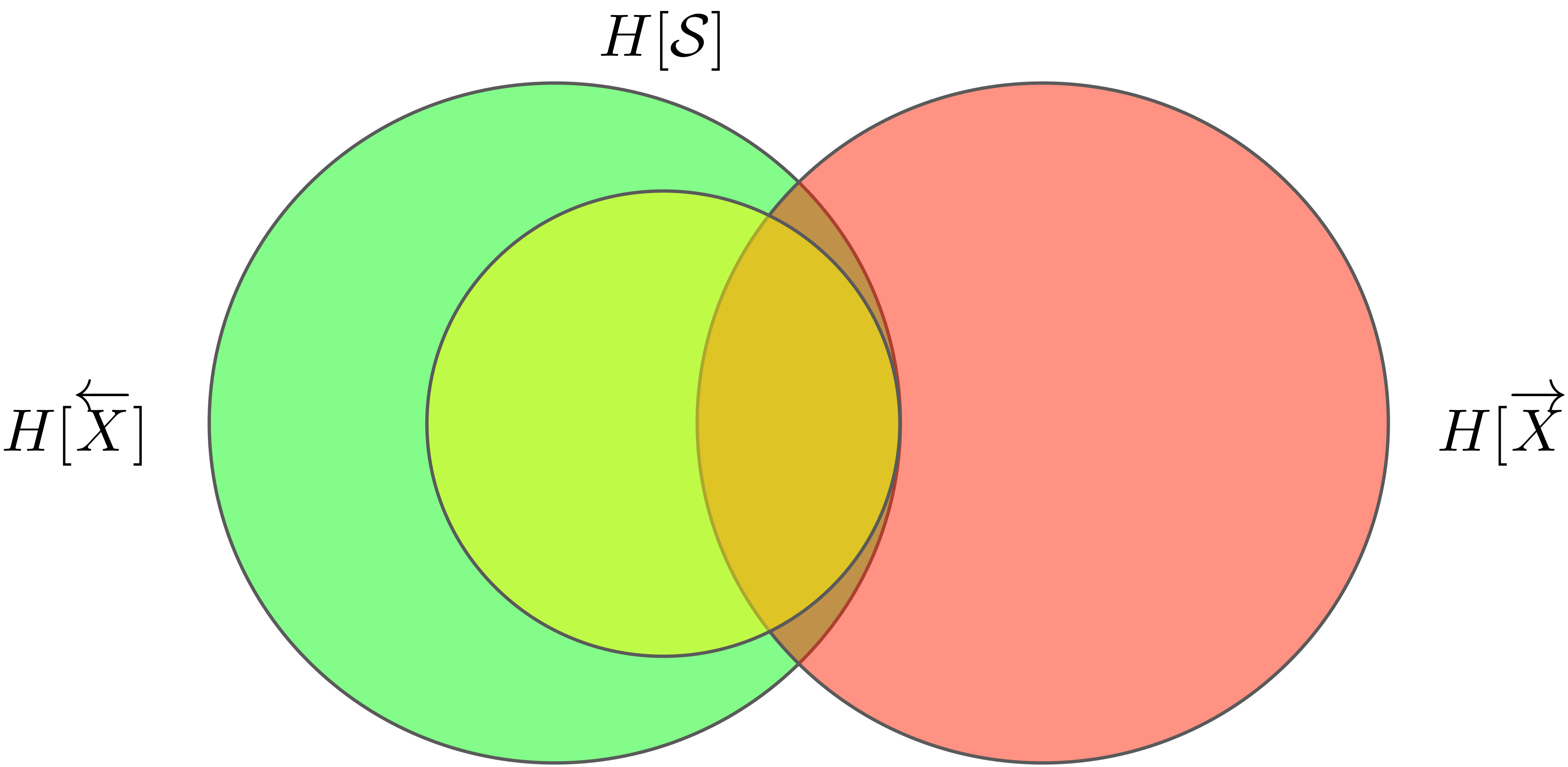
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



# Information Diagrams for Processes

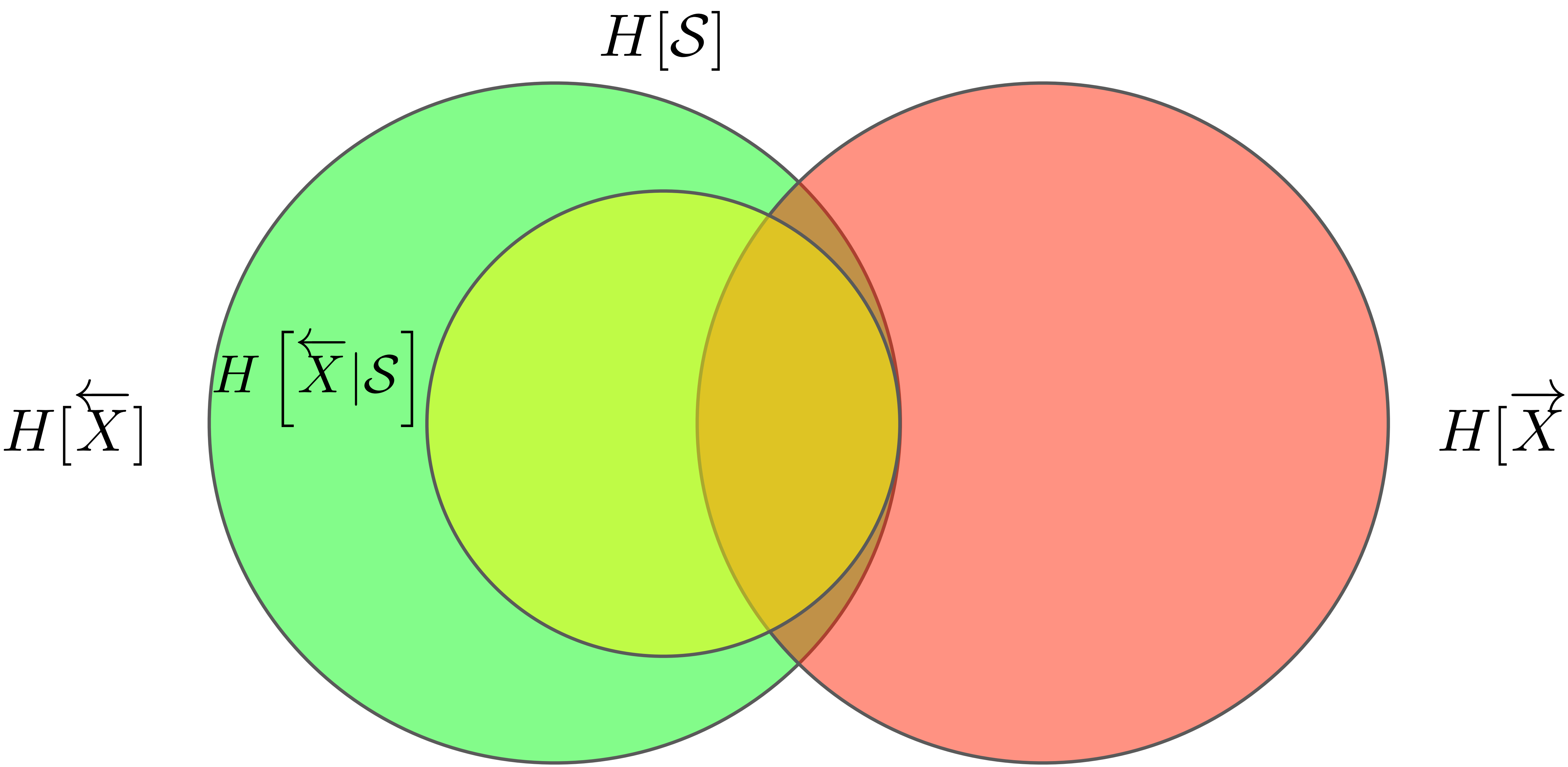
## Process I-diagram using ε-machine ...




Complexity Lecture 2: Computational Mechanics (CSSS 2011); Jim Crutchfield

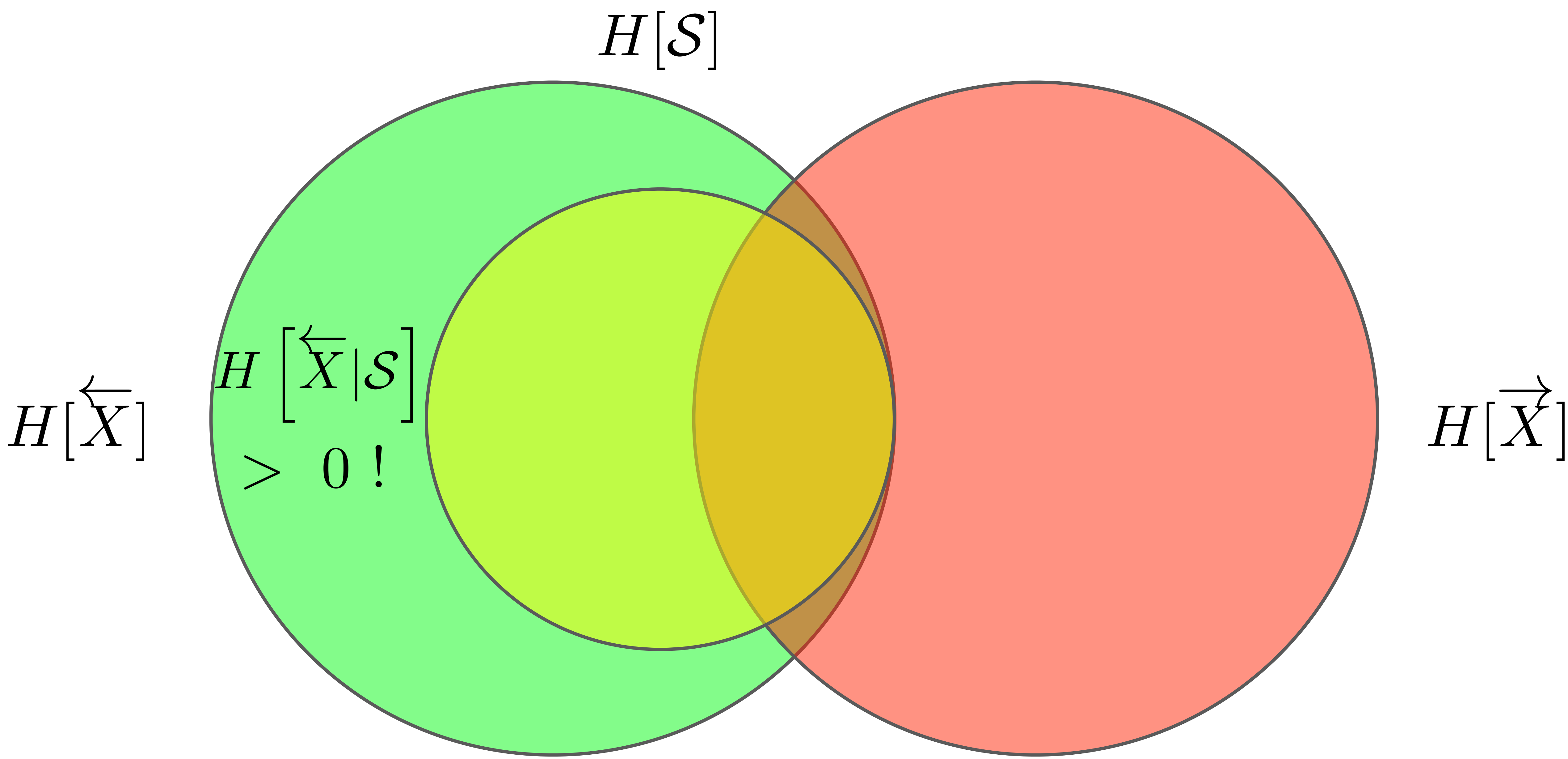
# Information Diagrams for Processes

## Process I-diagram using ε-machine ...



# Information Diagrams for Processes

## Process I-diagram using ε-machine ...

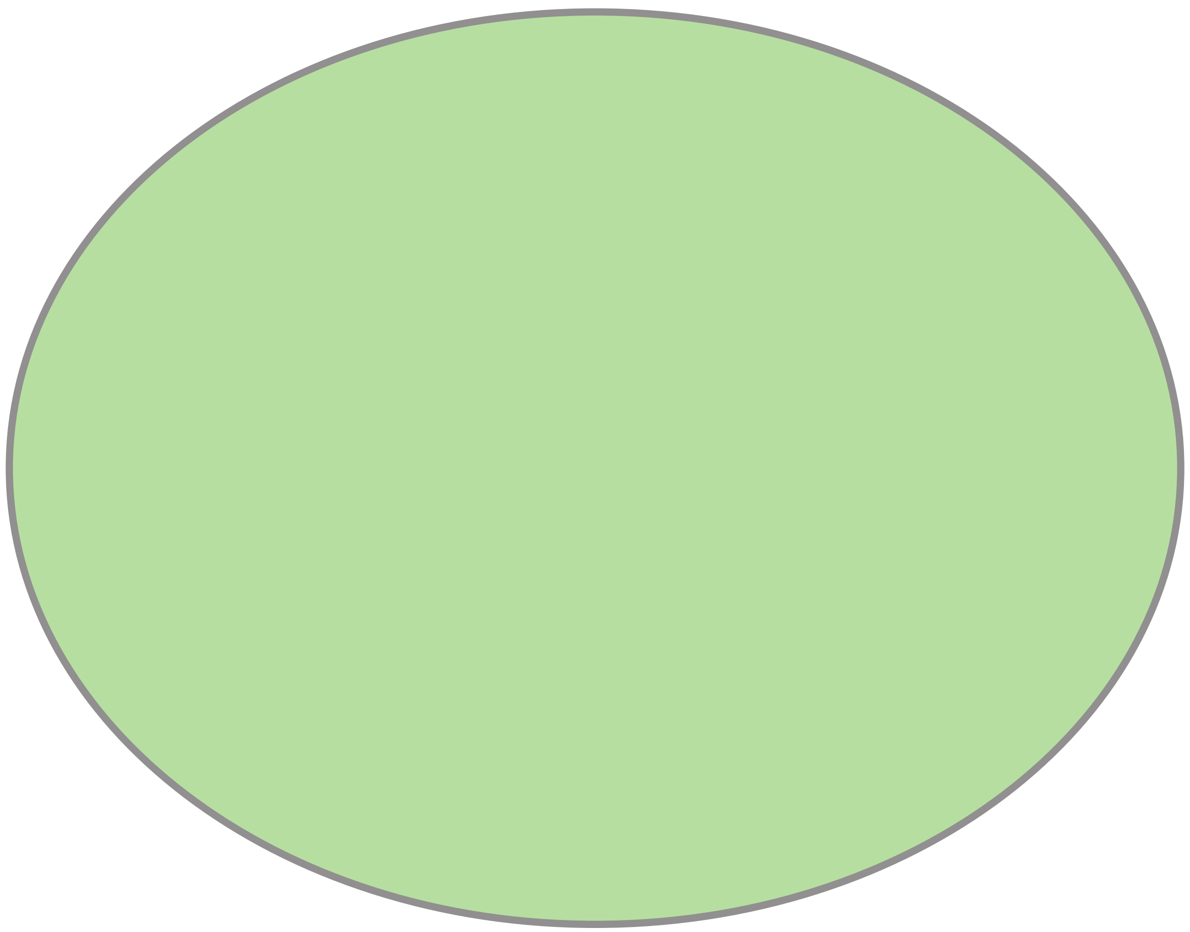


# Information Diagrams for Processes

## ε-Machine I-diagram:

# Information Diagrams for Processes

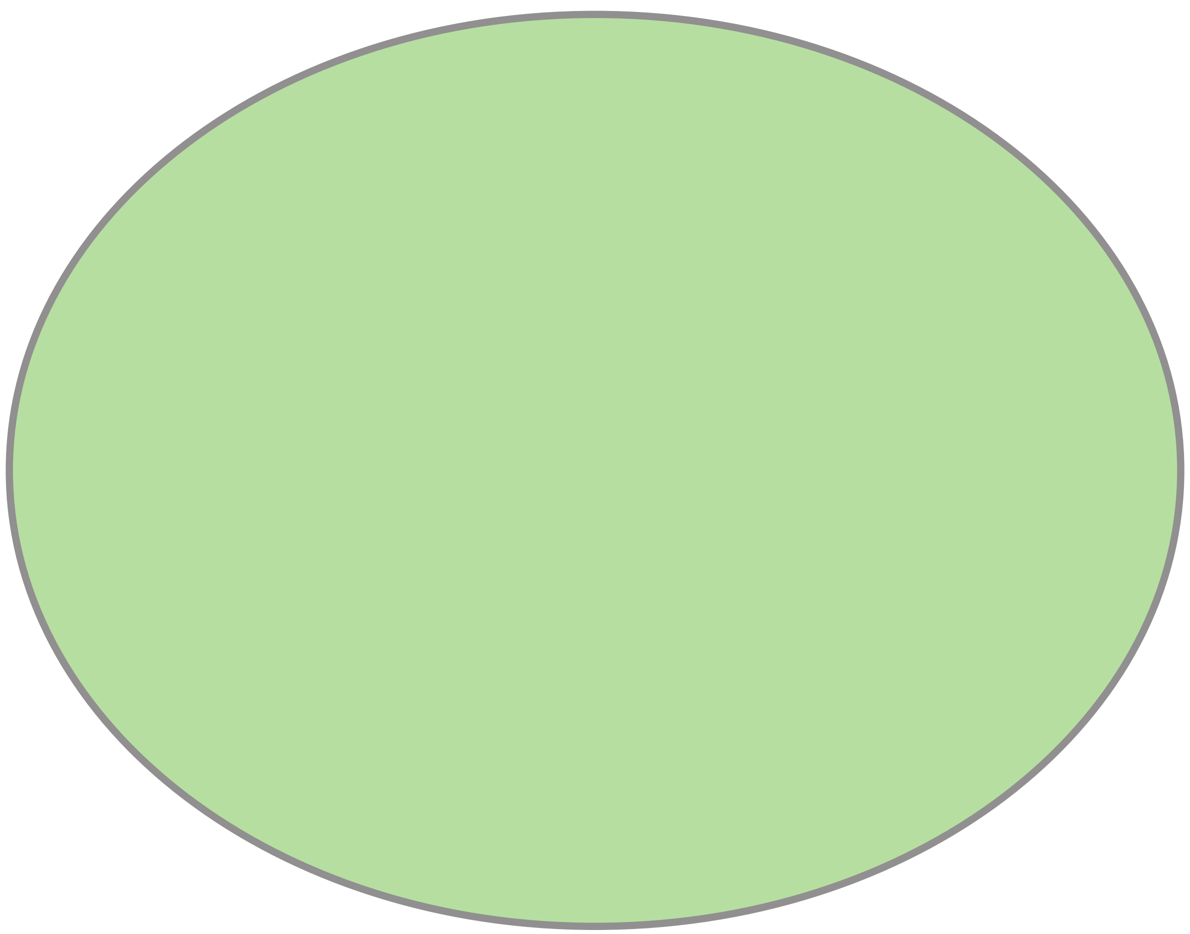
## ε-Machine I-diagram:



# Information Diagrams for Processes

## ε-Machine I-diagram:

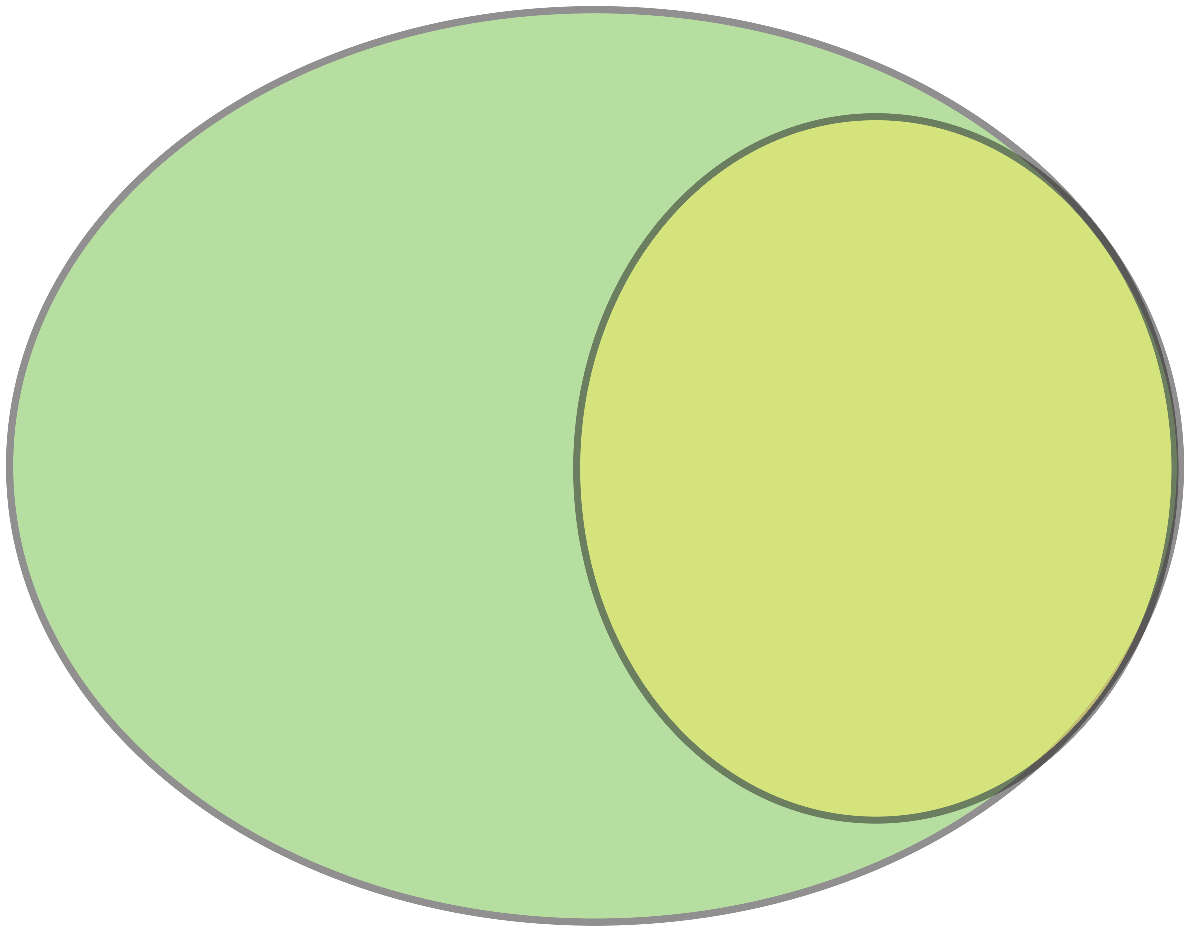
$H[\overleftarrow{X}]$



# Information Diagrams for Processes

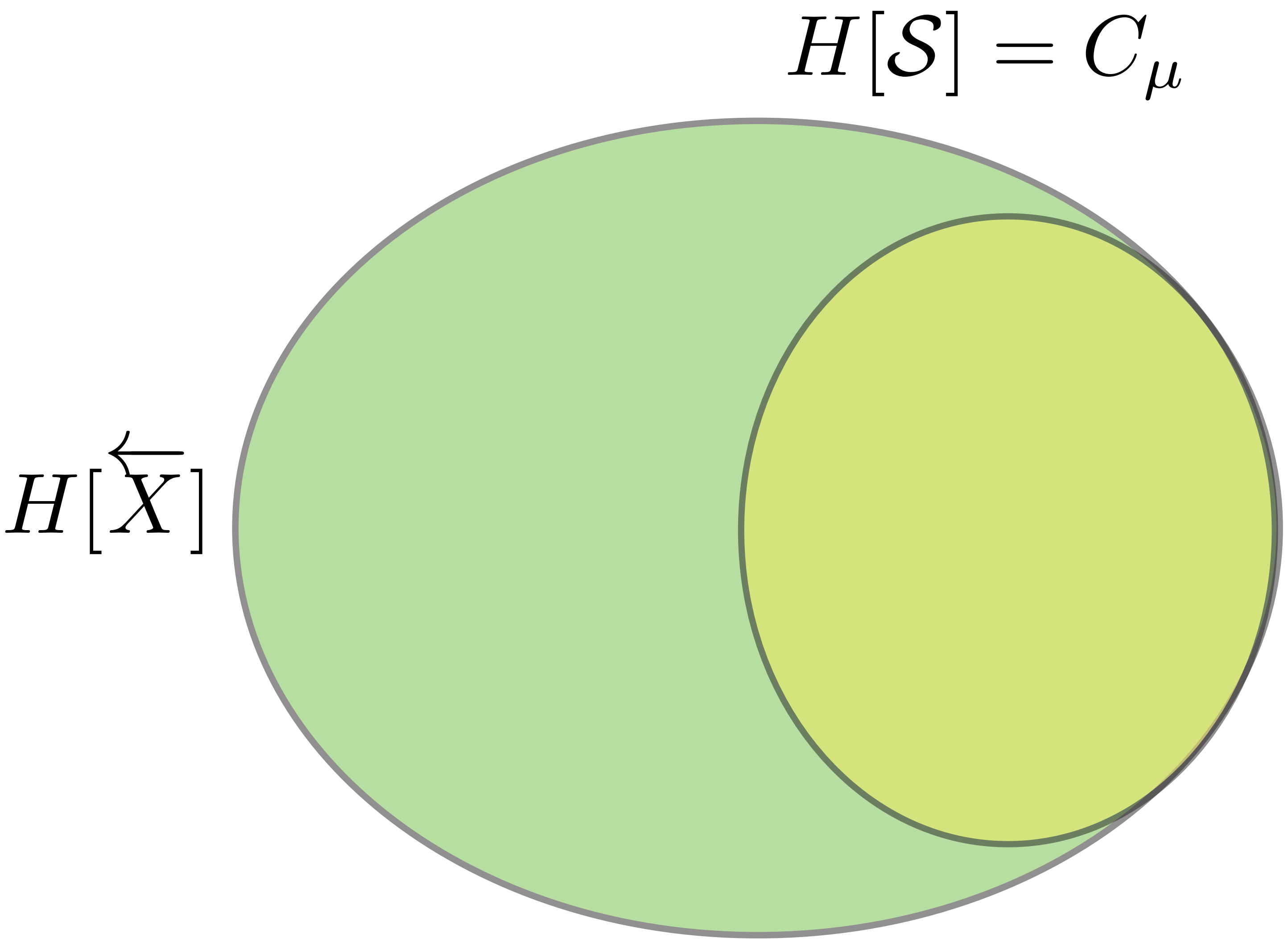
## ε-Machine I-diagram:

$H[\overleftarrow{X}]$



# Information Diagrams for Processes

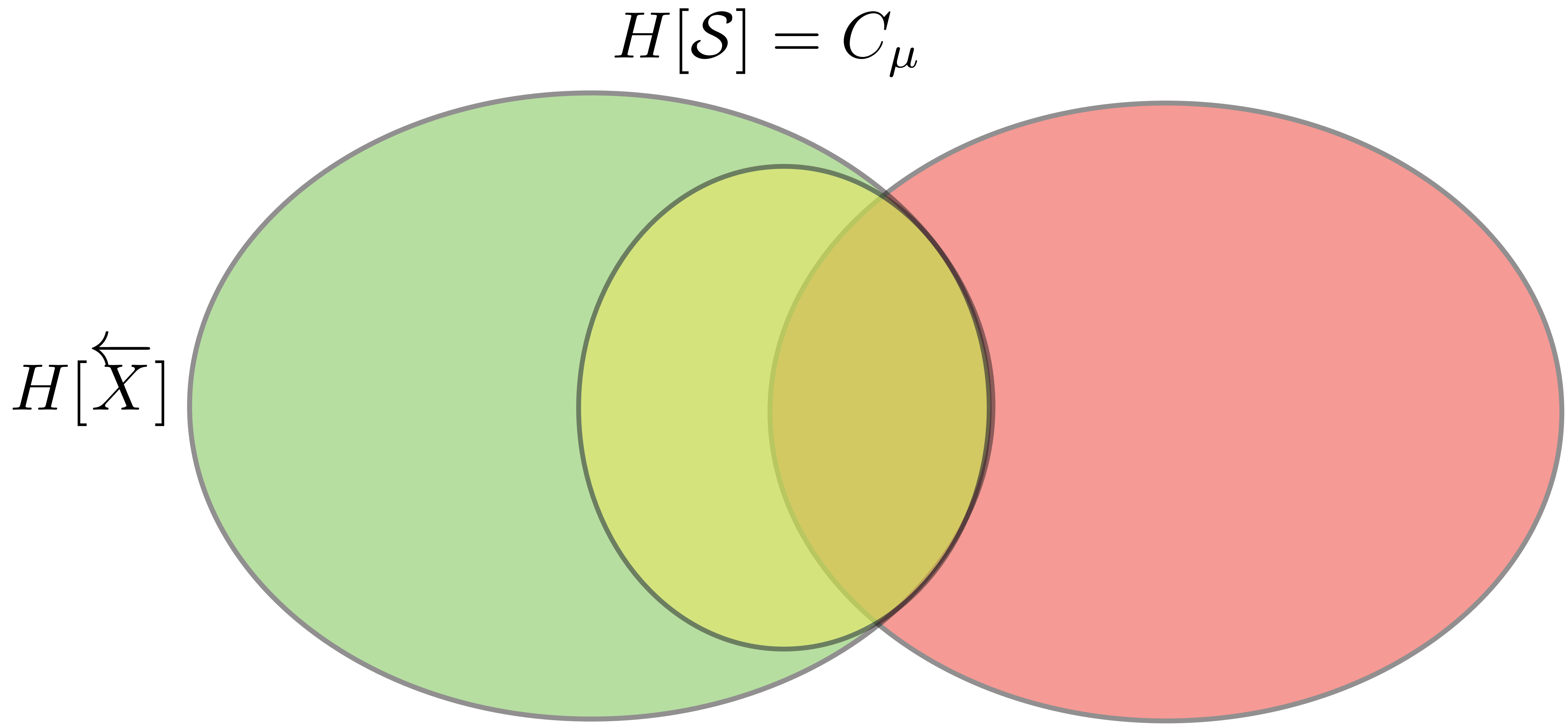
## ε-Machine I-diagram:




Complexity Lecture 2: Computational Mechanics (CSSS 2011); Jim Crutchfield

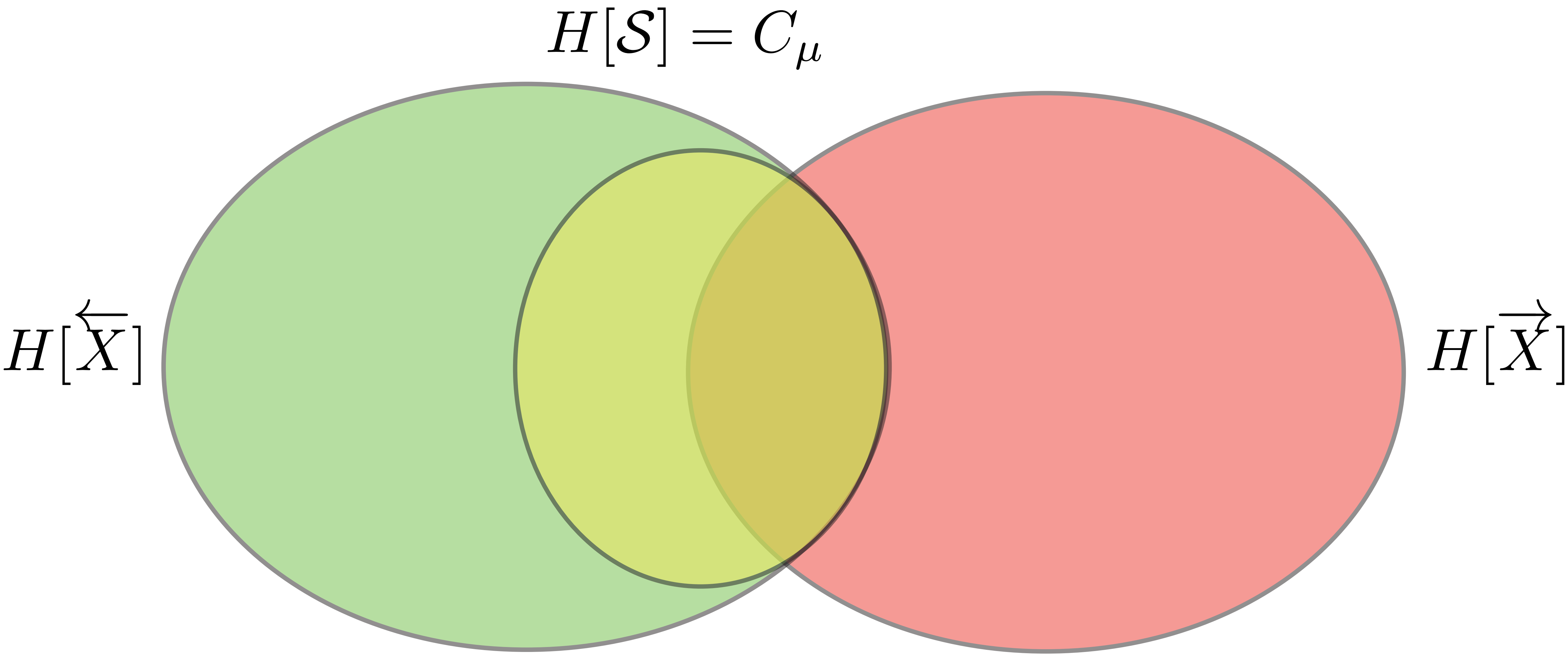
# Information Diagrams for Processes

## ε-Machine I-diagram:



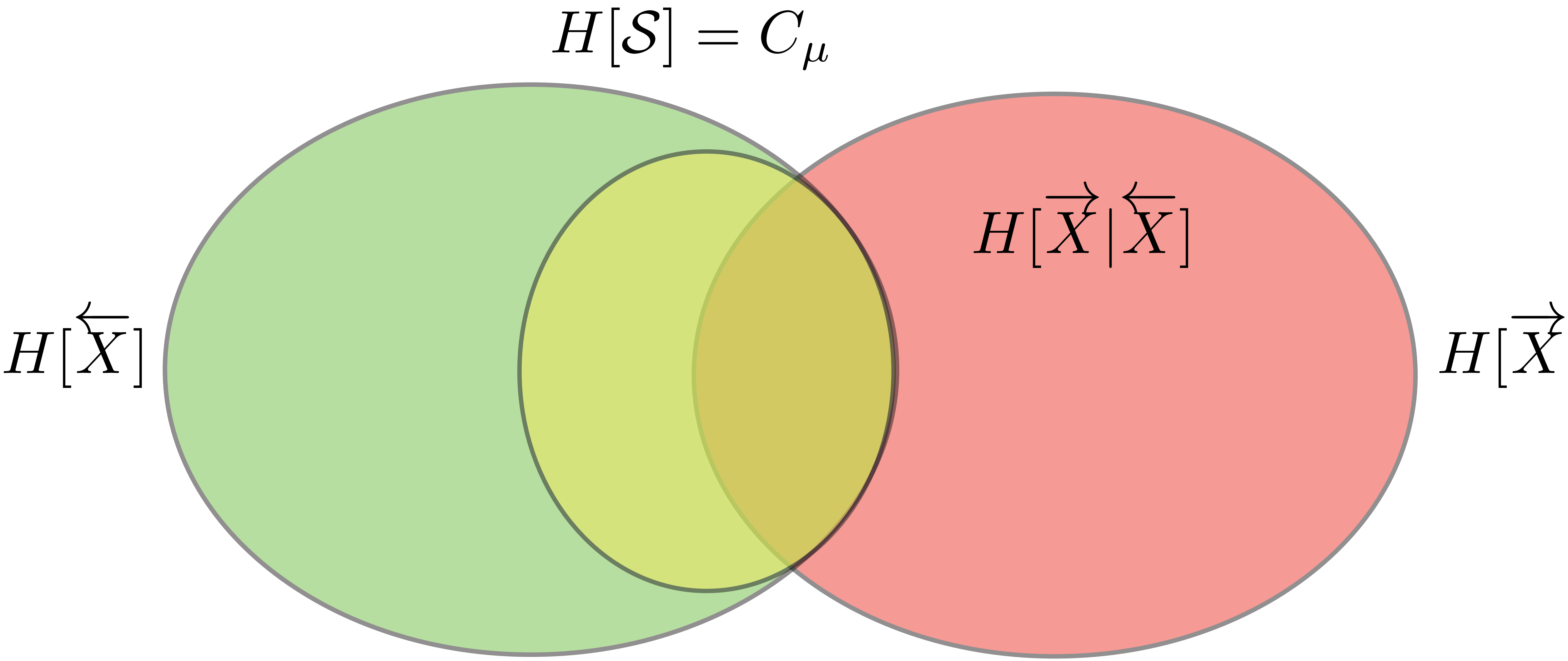
# Information Diagrams for Processes

## ε-Machine I-diagram:



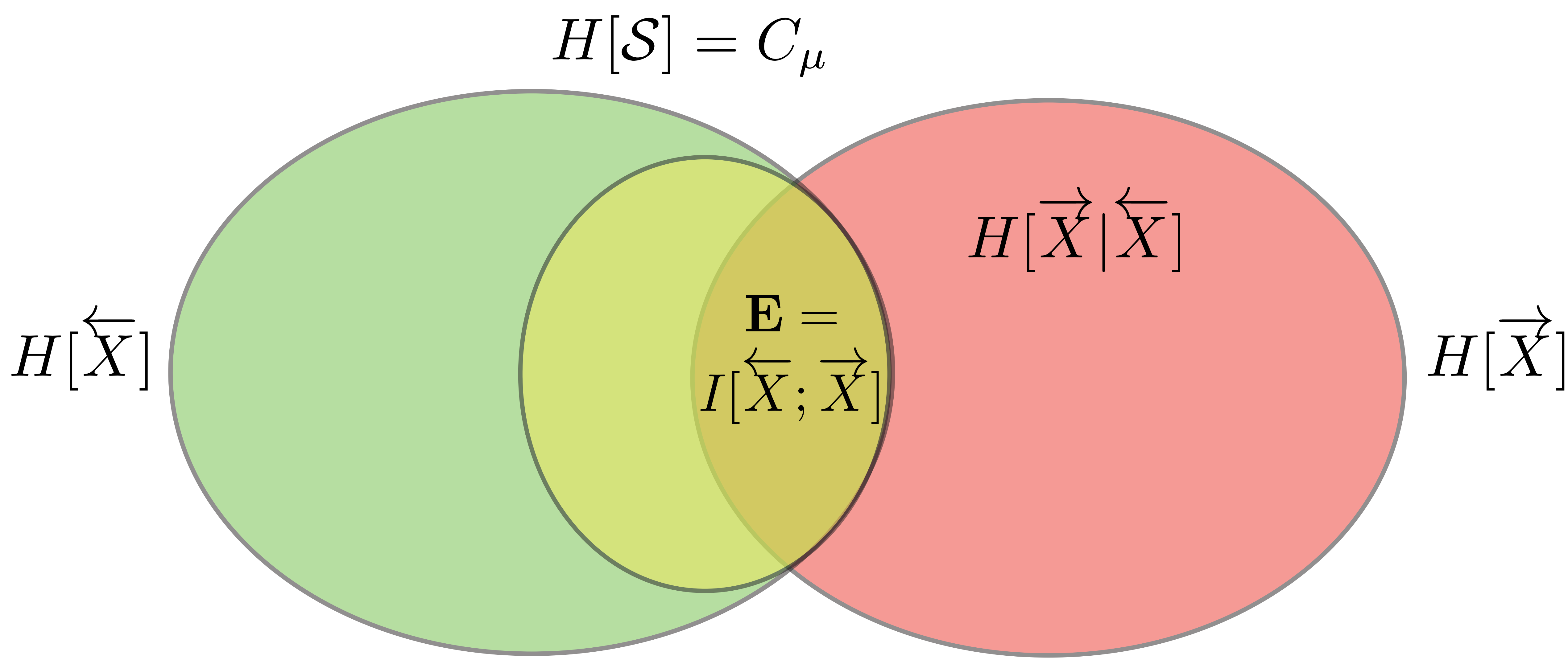
# Information Diagrams for Processes

## ε-Machine I-diagram:



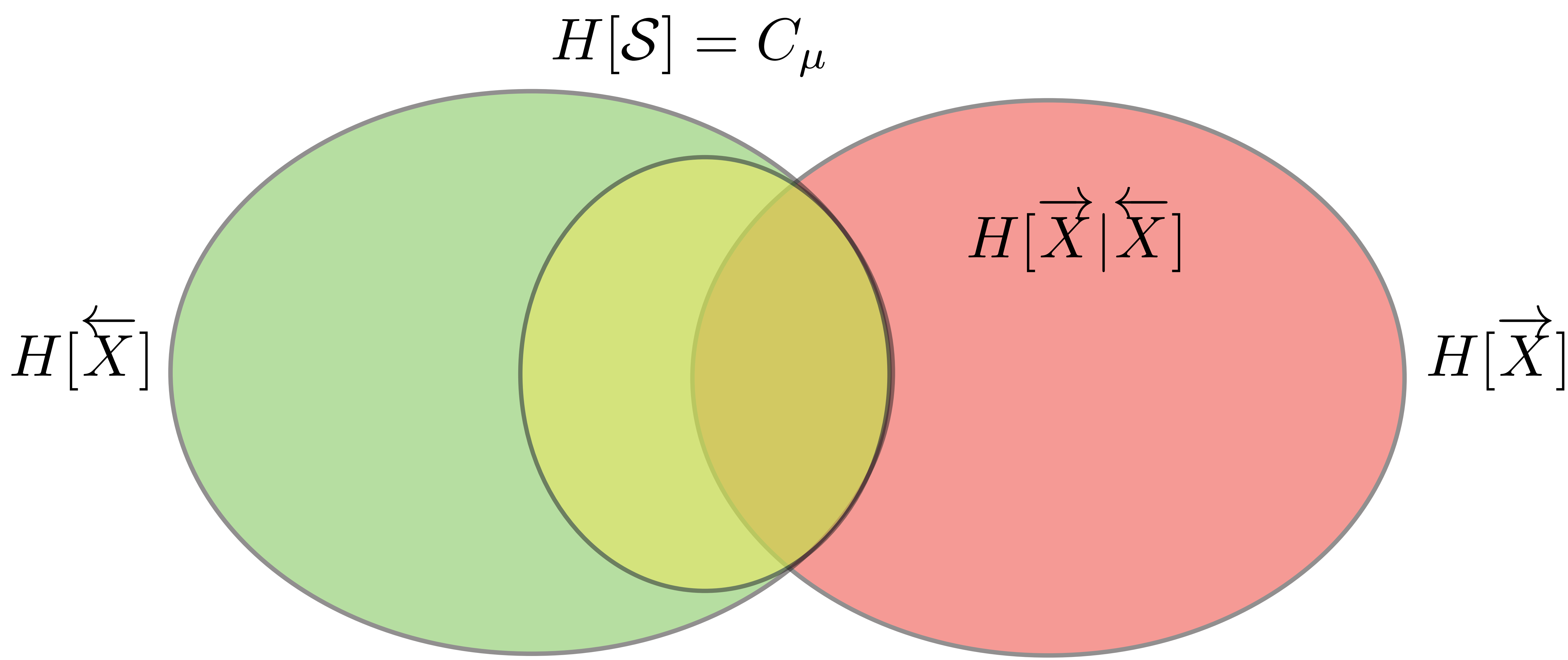
# Information Diagrams for Processes

## ε-Machine I-diagram:



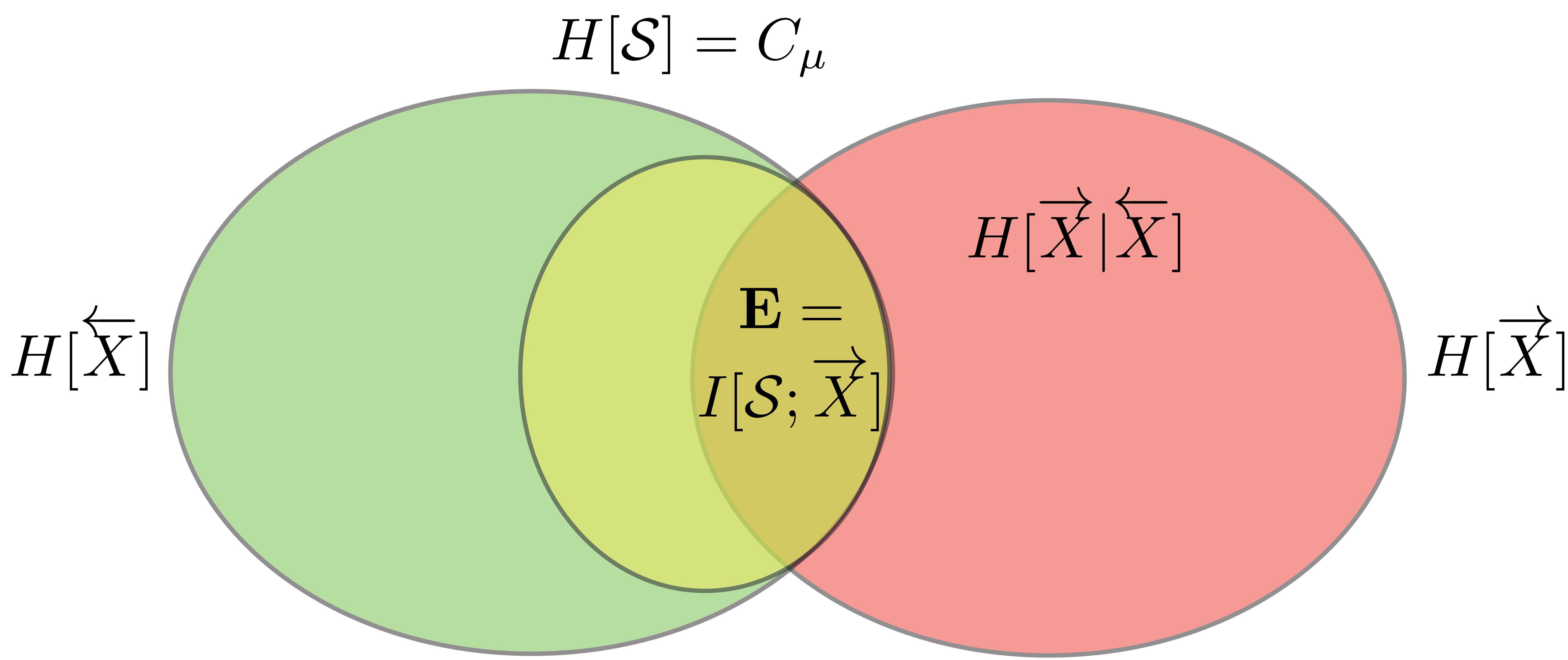
# Information Diagrams for Processes

## ε-Machine I-diagram:



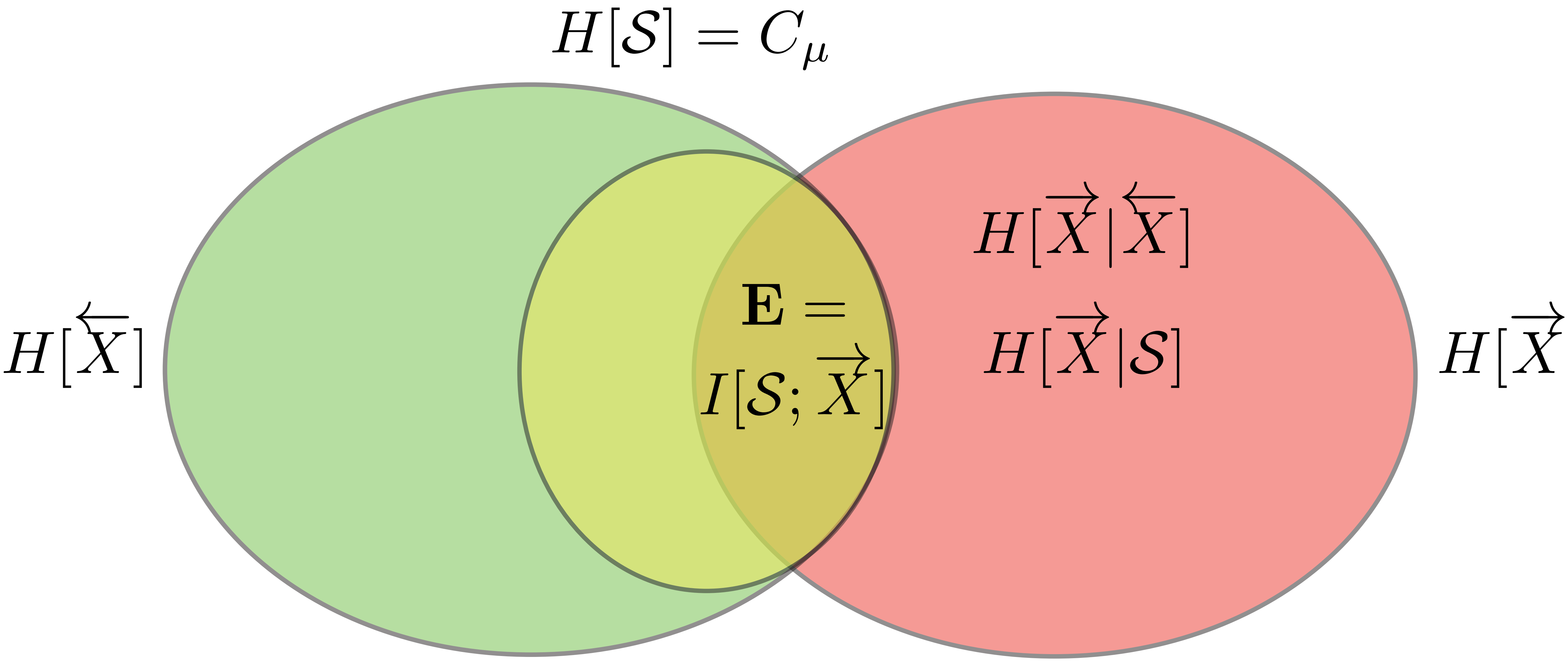
# Information Diagrams for Processes

## ε-Machine I-diagram:



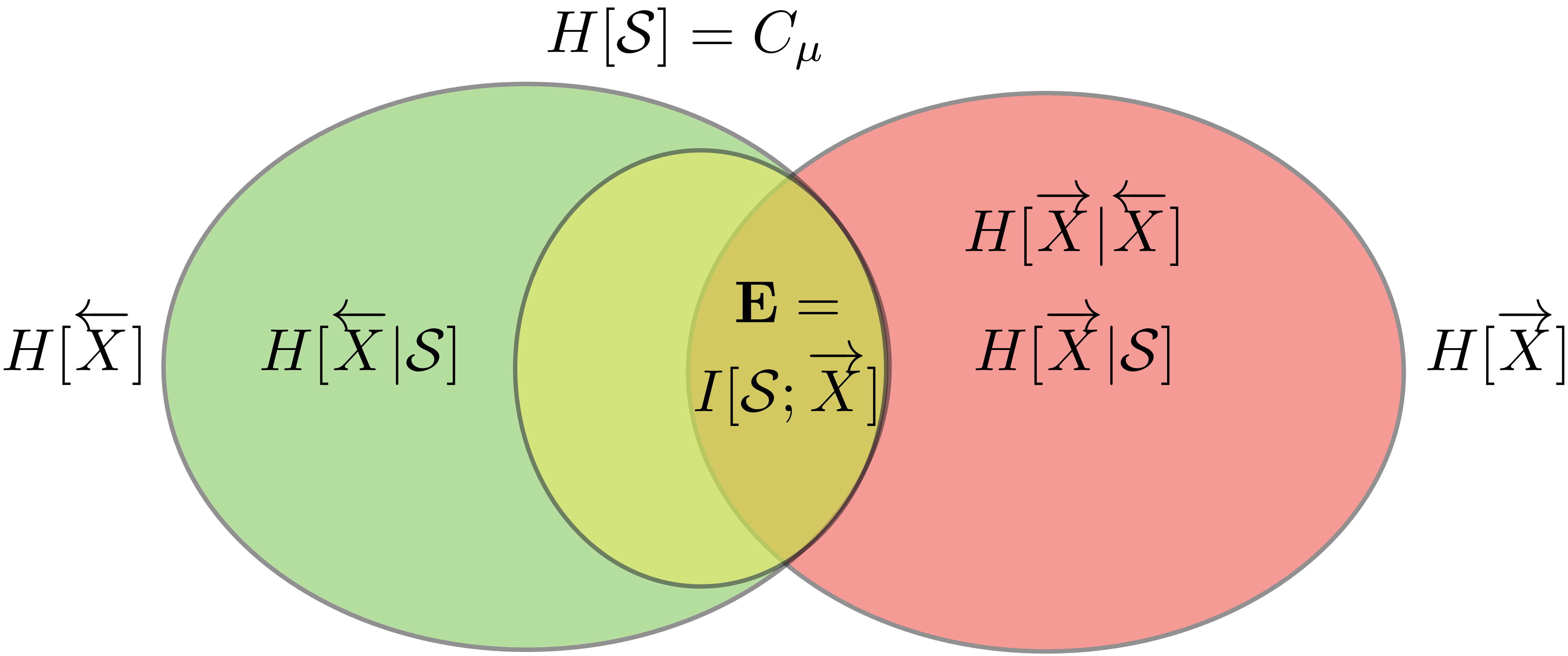
# Information Diagrams for Processes

## ε-Machine I-diagram:



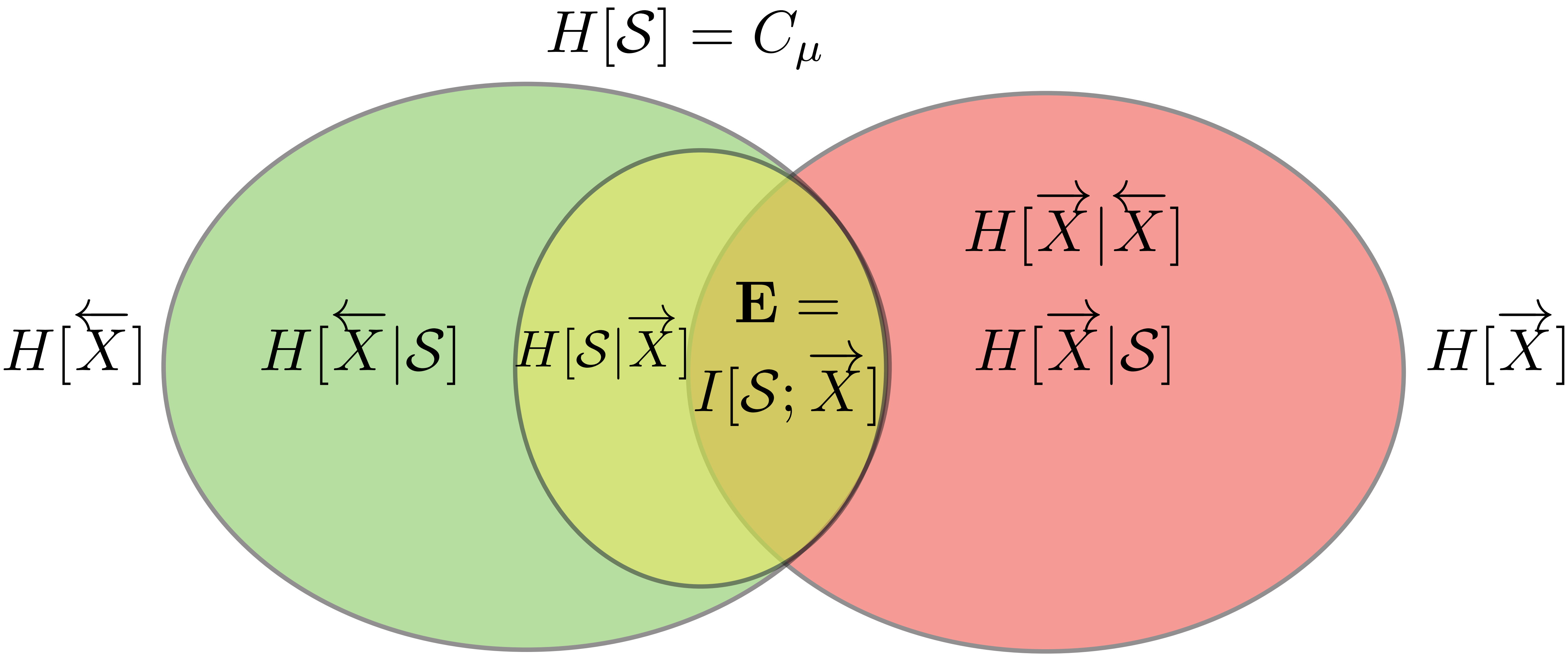
# Information Diagrams for Processes

## ε-Machine I-diagram:



# Information Diagrams for Processes

## ε-Machine I-diagram:



## Information Diagrams for Processes

What is $H[\overrightarrow{X}|\mathcal{S}]$?

Unpredictability: $H[\overrightarrow{X}^L|\mathcal{S}] = Lh_\mu$

Proof Sketch:

$$
\begin{aligned}
H[\overrightarrow{X}^L|\mathcal{S}] &= H[\overrightarrow{X}^L|\overleftarrow{X}] \\
&= H[X_0 X_1 \ldots X_{L-1}|\overleftarrow{X}] \\
&= H[X_1 \ldots X_{L-1}|\overleftarrow{X} X_0] + H[X_0|\overleftarrow{X}] \\
&= H[X_1 \ldots X_{L-1}|\overleftarrow{X}] + H[X_0|\overleftarrow{X}] \\
&\vdots \\
&= H[X_{L-1}|\overleftarrow{X}] + \cdots + H[X_1|\overleftarrow{X}] + H[X_0|\overleftarrow{X}] \\
&= LH[X_0|\overleftarrow{X}] \\
&= Lh_\mu
\end{aligned}
$$

# Information Diagrams for Processes

What is Mystery Wedge? $H[\mathcal{S}|\overrightarrow{X}]$

Uncertainty of causal state given future. Implications?

Recall Bound on Excess Entropy: $\mathbf{E} \leq C_\mu$

Proof sketch:
$$
\begin{aligned}
\mathbf{E} &= I[\overleftarrow{X};\overrightarrow{X}] \\
&= H[\overrightarrow{X}] - H[\overrightarrow{X}|\overleftarrow{X}] \\
&= H[\overrightarrow{X}] - H[\overrightarrow{X}|\mathcal{S}] \\
&= I[\overrightarrow{X};\mathcal{S}] \\
&= H[\mathcal{S}] - H[\mathcal{S}|\overrightarrow{X}] \\
&\leq H[\mathcal{S}] \\
&= C_\mu
\end{aligned}
$$

# Information Diagrams for Processes

What is Mystery Wedge? $H[\mathcal{S}|\overrightarrow{X}]$

Uncertainty of causal state given future. Implications?

Recall Bound on Excess Entropy: $\mathbf{E} \leq C_\mu$

Proof sketch:
$$\begin{aligned}
\mathbf{E} &= I[\overleftarrow{X};\overrightarrow{X}] \\
&= H[\overrightarrow{X}] - H[\overrightarrow{X}|\overleftarrow{X}] \\
&= H[\overrightarrow{X}] - H[\overrightarrow{X}|\mathcal{S}] \\
&= I[\overrightarrow{X};\mathcal{S}] \\
&= H[\mathcal{S}] - H[\mathcal{S}|\overrightarrow{X}] \\
&\leq H[\mathcal{S}] \\
&= C_\mu
\end{aligned}$$

I am the Mystery Wedge!

$\Box$

# Information Diagrams for Processes

What is Mystery Wedge? $H[\mathcal{S}|\overrightarrow{X}]$

Uncertainty of causal state given future. Implications?

Recall Bound on Excess Entropy: $\mathbf{E} \leq C_\mu$

Proof sketch:
$$
\begin{aligned}
\mathbf{E} &= I[\overleftarrow{X};\overrightarrow{X}] \\
&= H[\overrightarrow{X}] - H[\overrightarrow{X}|\overleftarrow{X}] \\
&= H[\overrightarrow{X}] - H[\overrightarrow{X}|\mathcal{S}] \\
&= I[\overrightarrow{X};\mathcal{S}] \\
&= H[\mathcal{S}] - H[\mathcal{S}|\overrightarrow{X}] \\
&\leq H[\mathcal{S}] \\
&= C_\mu
\end{aligned}
$$

I am the Mystery Wedge!

Wedge is the inaccessibility of hidden state information!

$H[\mathcal{S}|\overrightarrow{X}] = C_\mu - \mathbf{E}$ Wedge controls Internal - Observed

## What is Mystery Wedge? $H[\mathcal{S}|\overrightarrow{X}]$

Wedge is the inaccessibility of hidden state information!

$$H[\mathcal{S}|\overrightarrow{X}] = C_\mu - \mathbf{E}$$

The process crypticity:

$$\chi = C_\mu - \mathbf{E}$$

Controls how much internal state information is observable.

# Information Diagrams for Processes

RGJ:

- Markov order
- Cryptic order