The Past & the Future in the Present

Jim Crutchfield
Complexity Sciences Center
Physics Department
University of California at Davis

107th Statistical Mechanics Conference
Rutgers University

May 6-8, 2012

Joint work with Chris Ellison (UC Davis) & John Mahoney (UC Merced)

Information-Theoretic Analysis of Complex Systems

- Chain: $\overleftrightarrow{X} = \overleftarrow{X}_t \overrightarrow{X}_t$
- Past: $\overleftarrow{X}_t = \ldots X_{t-3} X_{t-2} X_{t-1}$
- Future: $\overrightarrow{X}_t = X_t X_{t+1} X_{t+2} \ldots$
- L-Block: $X_t^L = X_t X_{t+1} \ldots X_{t+L-1}$

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

Laplace's Spacetime Crystal

...

All moments, past, present, and future, always have existed, always will exist. The Tralfamadorians can look at all the different moments just the way we can look at a stretch of the Rocky Mountains, for instance. They can see how permanent all the moments are, and they can look at any moment that interests them. It is just an illusion we have here on Earth that one moment follows another one, like beads on a string, and that once a moment is gone it is gone forever.

Kurt Vonnegut, ***Slaughterhouse-Five*** (1968) p. 34.

Information-Theoretic Analysis of Complex Systems ...

- Process $\Pr(\overleftarrow{X}, \overrightarrow{X})$ is a communication channel from the past $\overleftarrow{X}$ to the future $\overrightarrow{X}$:

$$\text{Past} \longrightarrow \boxed{\text{Present}} \longrightarrow \text{Future}$$

Channel

Information-Theoretic Analysis of Complex Systems ...

- Process $\Pr(\overleftarrow{X}, \overrightarrow{X})$ is a communication channel from the past $\overleftarrow{X}$ to the future $\overrightarrow{X}$:

Past $\longrightarrow$ Present $\longrightarrow$ Future

Information Rate h_μ

Channel Capacity C

Information-Theoretic Analysis of Complex Systems ...

- Process $\Pr(\overleftarrow{X}, \overrightarrow{X})$ is a communication channel from the past $\overleftarrow{X}$ to the future $\overrightarrow{X}$:

Past $\longrightarrow$ Present $\longrightarrow$ Future

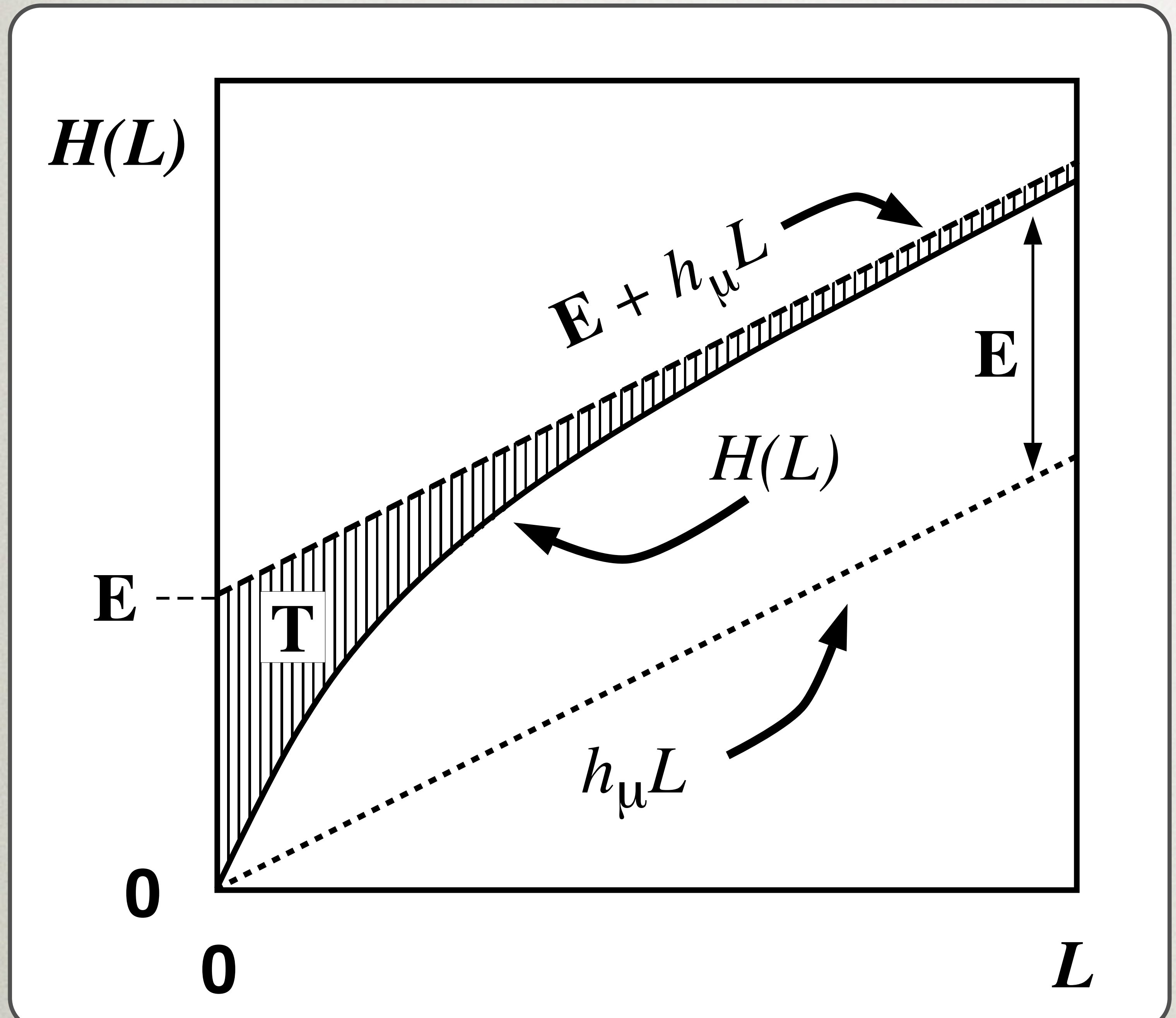
Information Rate h_μ

Channel Capacity C

- Channel Utilization: **Excess Entropy**

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

Roadmap to Information(s)



Block Entropy

$$H(L) = H[\Pr(X^L)]$$

J. P. Crutchfield and D. P. Feldman, "Regularities Unseen, Randomness Observed: Levels of Entropy Convergence", CHAOS **13**:1 (2003) 25-54.

Is Information Theory Sufficient?

- No!
- Measurements = process states? Wrong!
- Hidden processes
- No direct measure of structure

Computational Mechanics:
What are the hidden states?

- Group all histories that give same prediction:

$$\epsilon(\overleftarrow{x}) = \{\overleftarrow{x}' : \Pr(\overrightarrow{X}|\overleftarrow{x}) = \Pr(\overrightarrow{X}|\overleftarrow{x}')\}$$

- Equivalence relation: $\overleftarrow{x} \sim \overleftarrow{x}'$
- Equivalence classes are process's **causal states**:

$$\boldsymbol{\mathcal{S}} = \Pr(\overleftarrow{X}, \overrightarrow{X}) / \sim$$

- **ε-Machine**: Optimal, minimal, unique predictor.

J. P. Crutchfield, K. Young, "Inferring Statistical Complexity", Physical Review Letters **63** (1989) 105-108.

Computational Mechanics

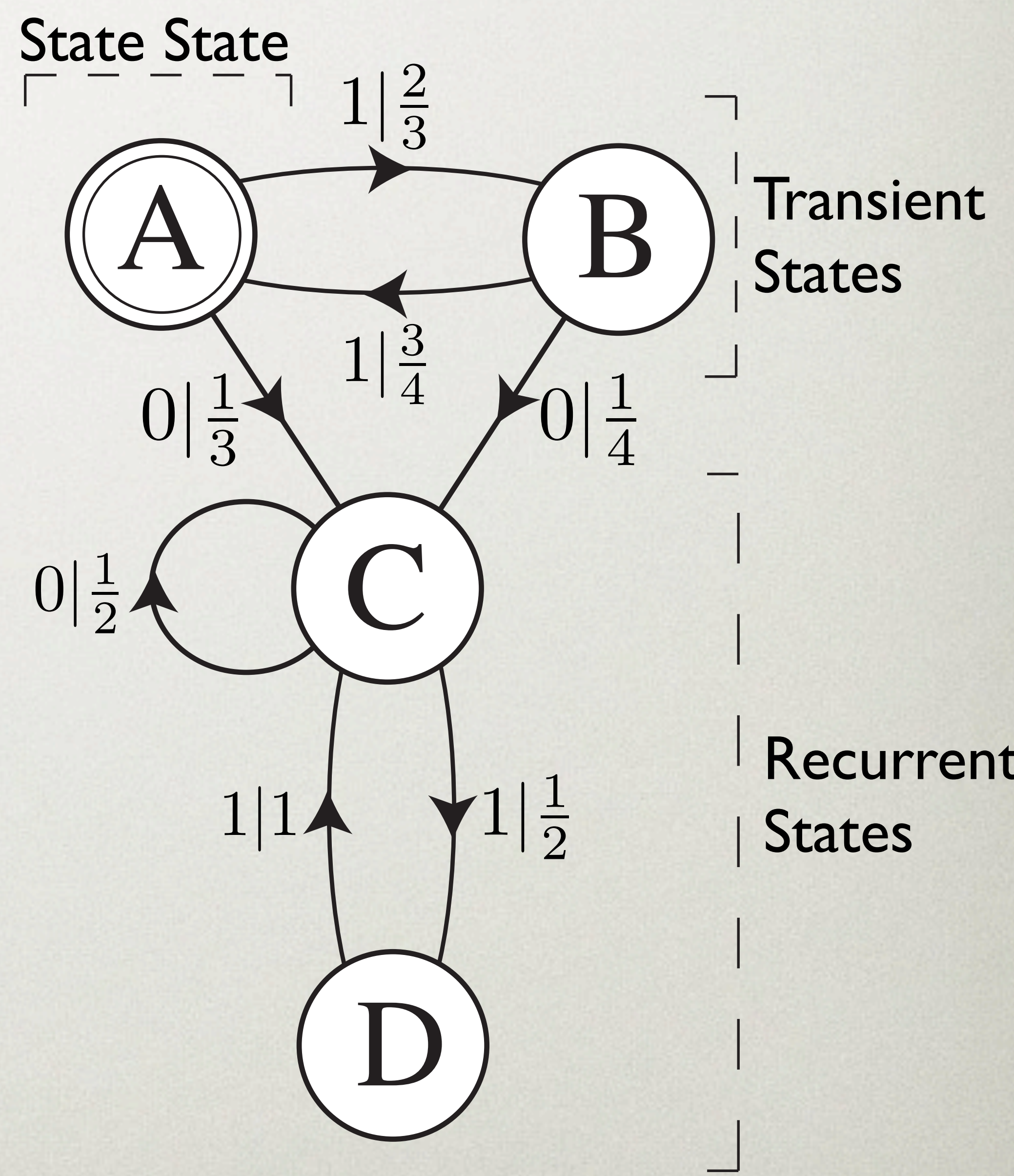
- ε-Machine:

$$M = \left\{ \boldsymbol{\mathcal{S}}, \{ T^{(x)} : x \in \mathcal{A} \} \right\}$$

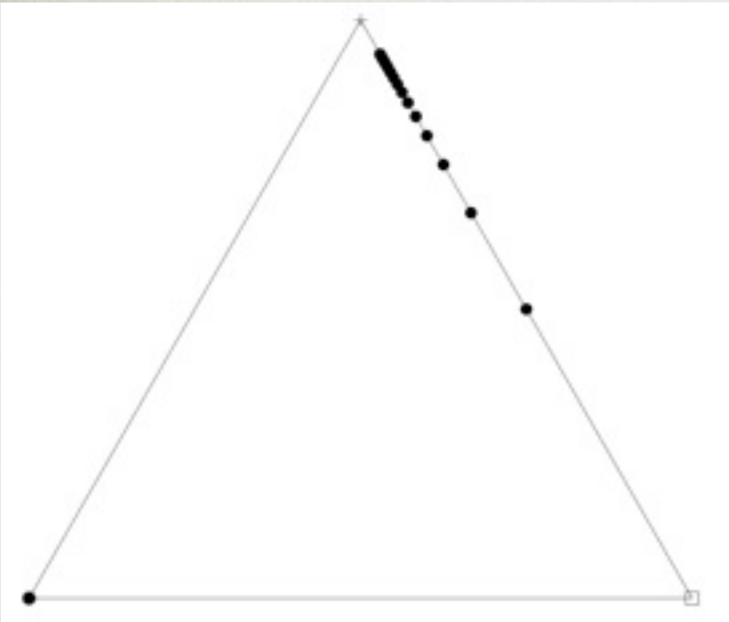
- Dynamic:

$$T^{(x)}_{\sigma,\sigma'} = \Pr(\sigma' | \sigma, x)$$

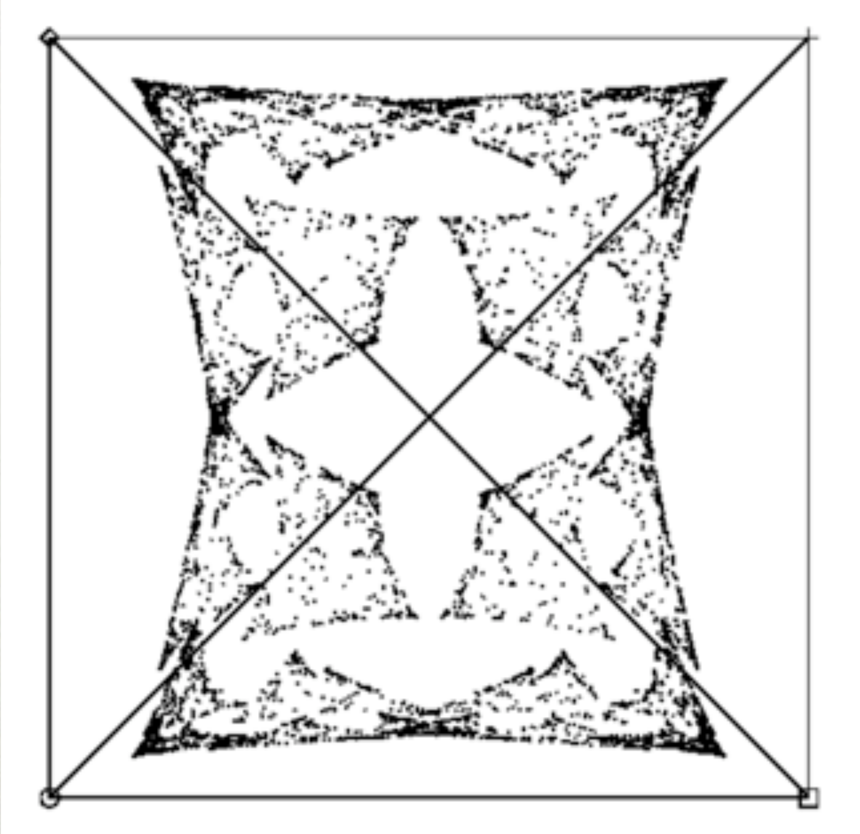
$$\sigma, \sigma' \in \boldsymbol{\mathcal{S}}$$



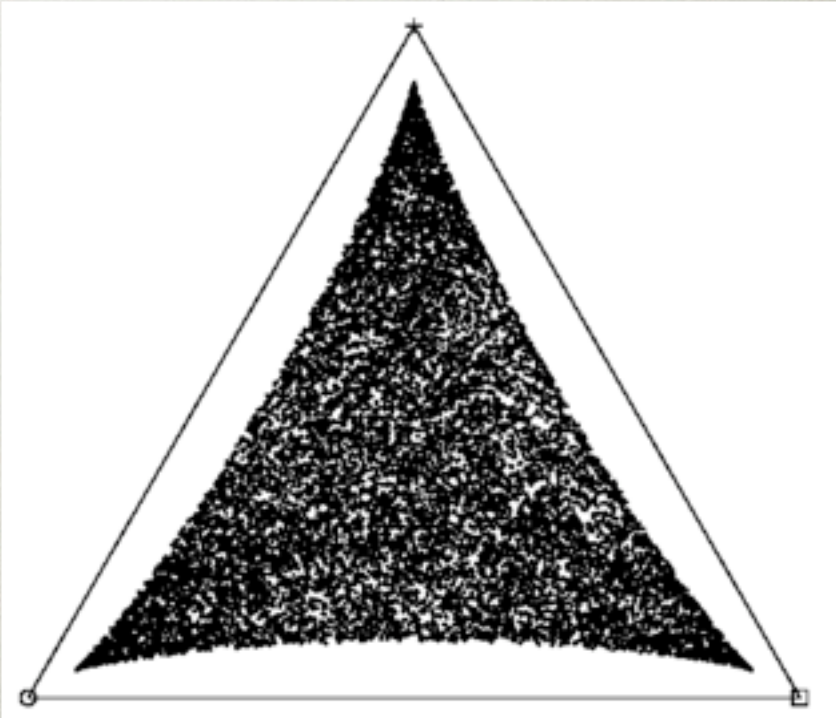
Varieties of ε-Machine



Denumerable Causal States



Fractal



Continuous

J. P. Crutchfield, "Calculi of Emergence: Computation, Dynamics, and Induction", Physica D **75** (1994) 11-54.

Kinds of Intrinsic Computing

- Directly from ε-Machine:
  - Stored information (**Statistical complexity**):

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

  - Information production (**Entropy rate**):

$$h_\mu = - \sum_{\sigma \in \boldsymbol{\mathcal{S}}} \Pr(\sigma) \sum_{\sigma' \in \boldsymbol{\mathcal{S}}, s \in \mathcal{A}} \Pr(\sigma \rightarrow_s \sigma') \log_2 \Pr(\sigma \rightarrow_s \sigma')$$

Prediction V. Modeling

- Hidden: State information via measurement.
- So, how accessible is information?
- How do measurements reveal internal states?
- Quantitative version:
  - Prediction ~ $\mathbf{E}$
  - Modeling ~ C_μ
- Can get h_μ and C_μ directly from ε-Machine.
- How to calculate $\mathbf{E}$ from ε-Machine?

Directional Computational Mechanics

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

  $\xrightarrow{\hspace{6em}}$

  Scan direction

- "Forward" ε-Machine: M^+
  - Equivalence Relation $\overrightarrow{x} \sim^+ \overrightarrow{x}'$: $\epsilon^+(\overrightarrow{x})$
  - Forward Causal States: $\boldsymbol{\mathcal{S}}^+$
  - Measures:
    - Entropy Rate: h_μ^+
    - Statistical Complexity: C_μ^+

Directional Computational Mechanics

- Now, **reverse ε-Machine**:

$$\overleftrightarrow{X} = \ldots X_{-2} X_{-1} X_0 X_1 X_2 \ldots$$

Scan direction

- Retrodictive equivalence relation: $\overrightarrow{x} \sim^- \overrightarrow{x}'$

$$\epsilon^-(\overrightarrow{x}) = \{\overrightarrow{x}' : \Pr(\overleftarrow{X}|\overrightarrow{x}) = \Pr(\overleftarrow{X}|\overrightarrow{x}')\}$$

- Retrodictive causal states: $\boldsymbol{\mathcal{S}}^- = \Pr(\overleftarrow{X}, \overrightarrow{X}) / \sim^-$
- Reverse ε-Machine: M^-
- Retrodictive entropy rate: h_μ^-
- Reverse statistical complexity: $C_\mu^- \equiv H[\mathcal{S}^-]$

Directional Computational Mechanics

- In which time direction most predictable?
- Excess entropy:
- Stored information?

J. P. Crutchfield, "Semantics and Thermodynamics", in **Nonlinear Modeling and Forecasting**, M. Casdagli and S. Eubank, editors, Addison-Wesley, Reading, Massachusetts (1992) 317-359.

Directional Computational Mechanics

- In which time direction most predictable?

  Neither! $h_\mu^- = h_\mu^+$

- Excess entropy:

- Stored information?

J. P. Crutchfield, "Semantics and Thermodynamics", in **Nonlinear Modeling and Forecasting**, M. Casdagli and S. Eubank, editors, Addison-Wesley, Reading, Massachusetts (1992) 317-359.

Directional Computational Mechanics

- In which time direction most predictable?

  Neither! $h_\mu^- = h_\mu^+$

- Excess entropy:

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

- Stored information?

J. P. Crutchfield, "Semantics and Thermodynamics", in **Nonlinear Modeling and Forecasting**, M. Casdagli and S. Eubank, editors, Addison-Wesley, Reading, Massachusetts (1992) 317-359.

Directional Computational Mechanics

- In which time direction most predictable?

Neither! $h_\mu^- = h_\mu^+$

- Excess entropy:

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

- Stored information?

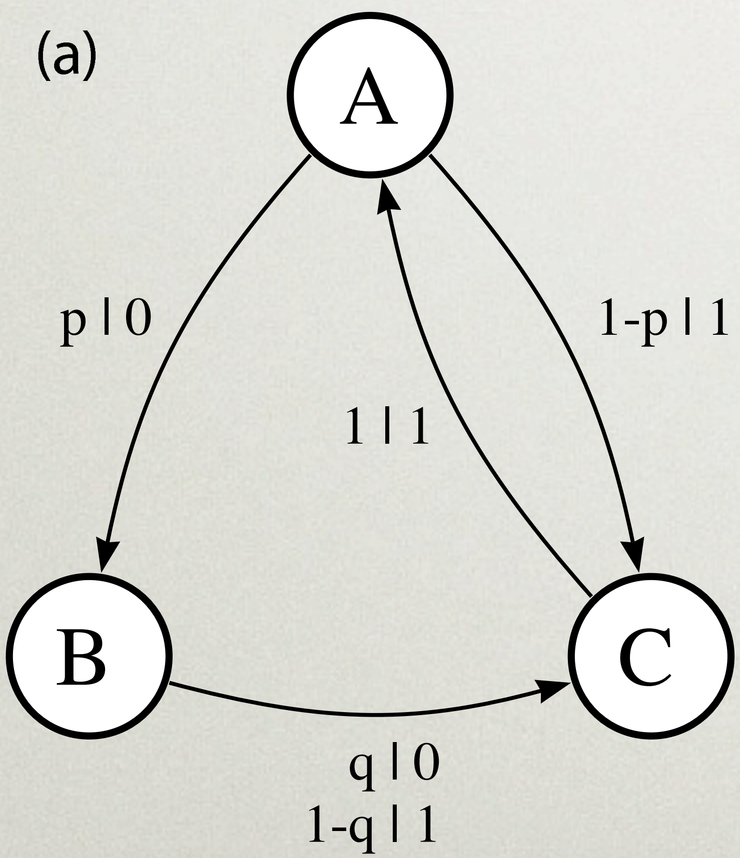
$$C_\mu^- \neq C_\mu^+$$

J. P. Crutchfield, "Semantics and Thermodynamics", in **Nonlinear Modeling and Forecasting**, M. Casdagli and S. Eubank, editors, Addison-Wesley, Reading, Massachusetts (1992) 317-359.

Directional Computational Mechanics

- Random Insertion Process

Forward ε-machine

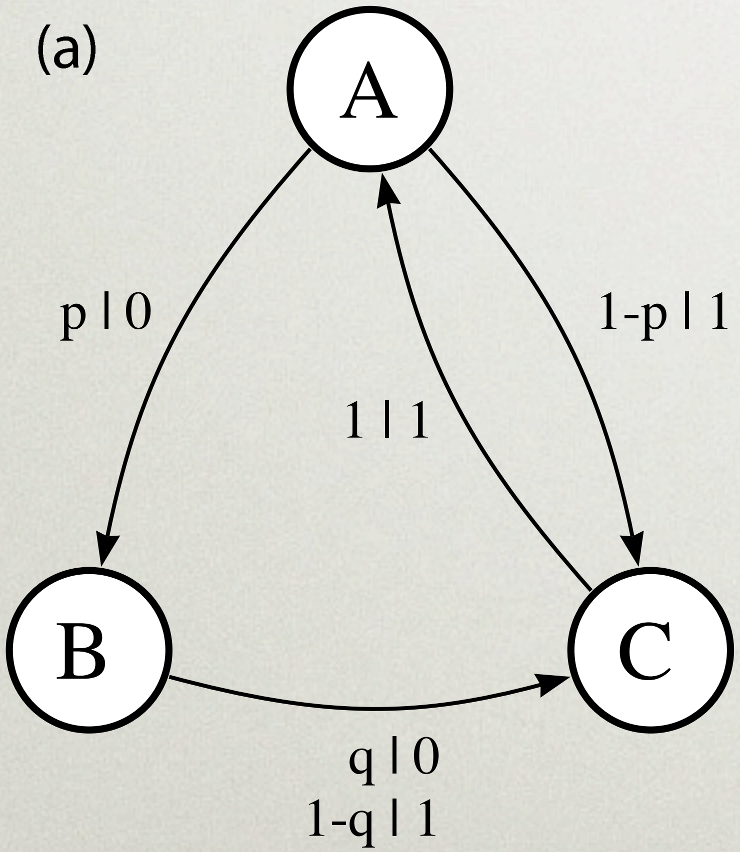


Directional Computational Mechanics

- Random Insertion Process

Forward ε-machine

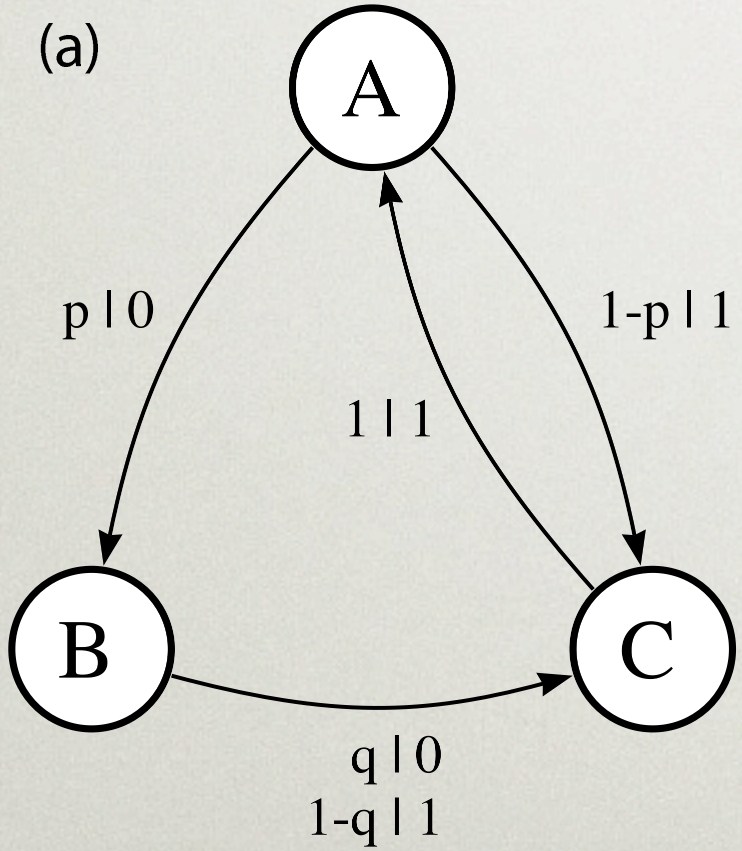
Reverse ε-machine



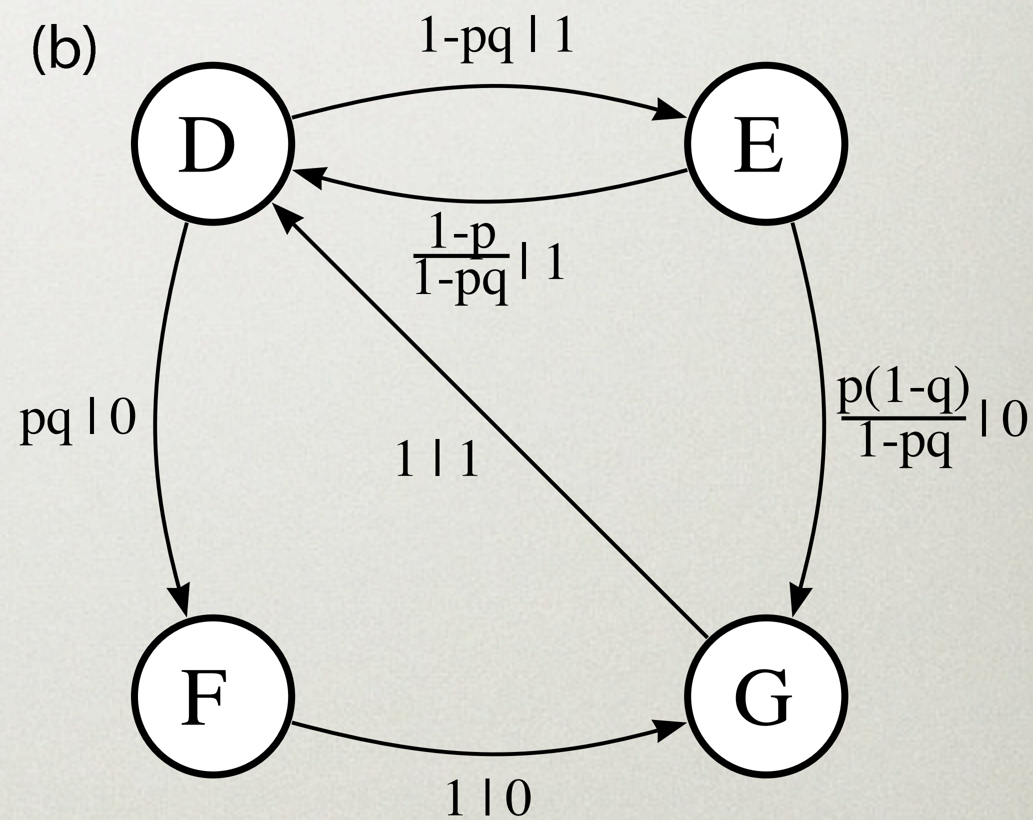
Directional Computational Mechanics

- Random Insertion Process

Forward ε-machine



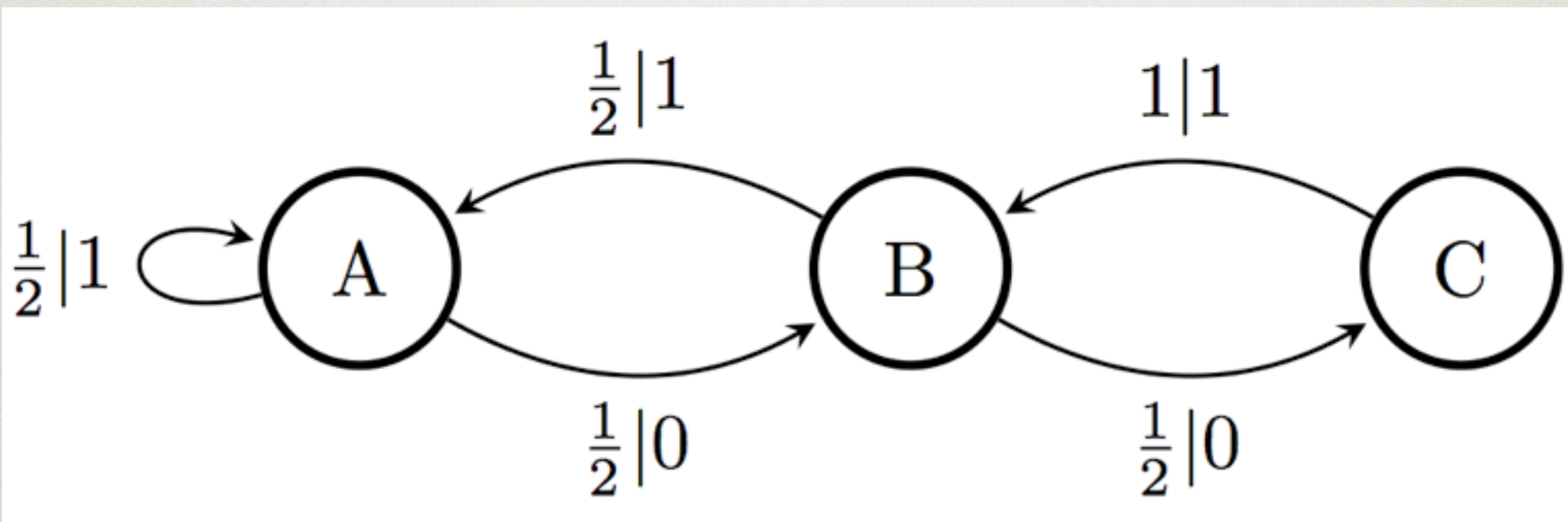
Reverse ε-machine



Directional Computational Mechanics

- At Most Two 0s + Isolated 1 $\Rightarrow$ at most One 0

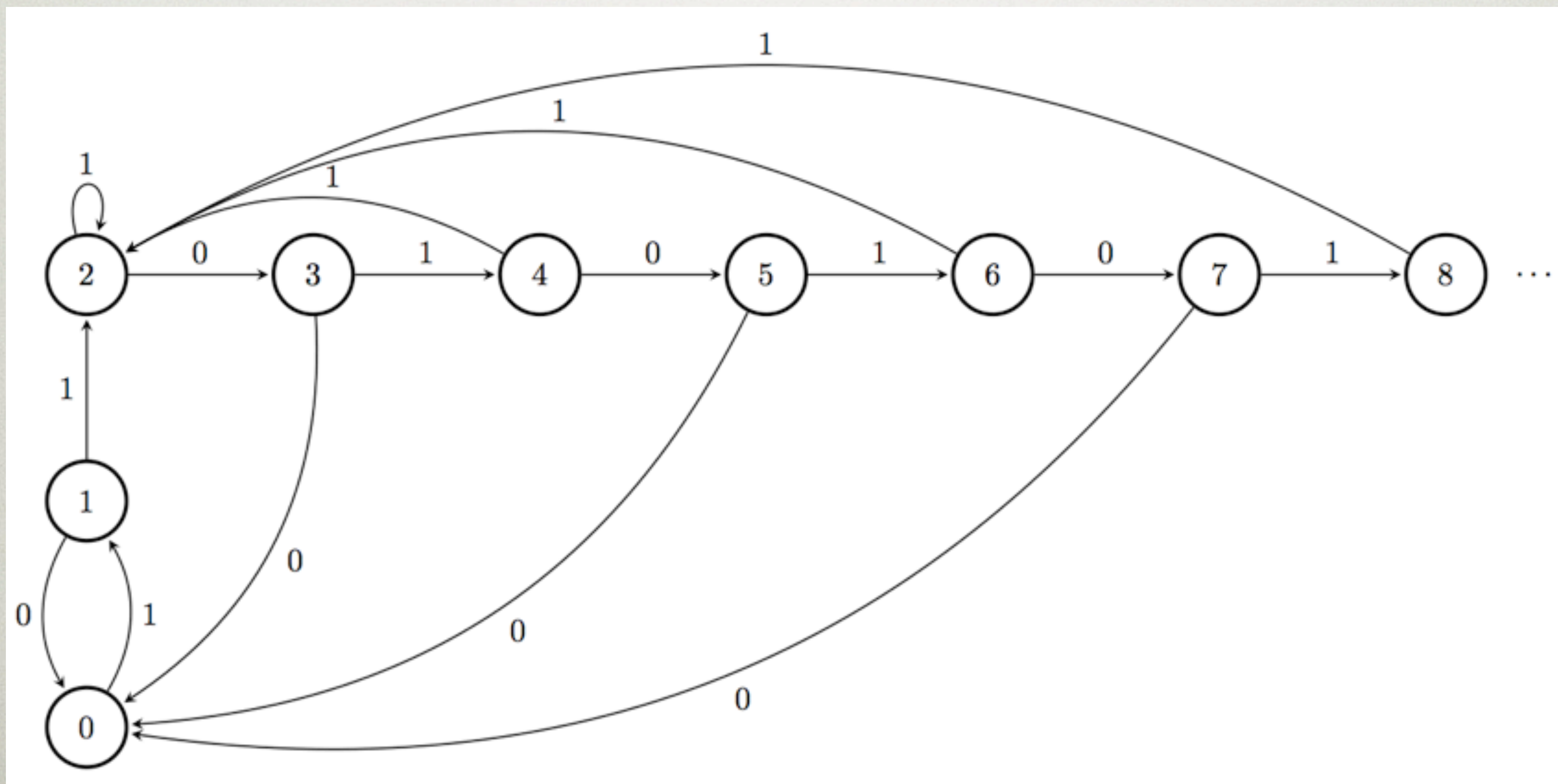
Forward ε-machine



Directional Computational Mechanics

- At Most Two 0s + Isolated 1 ⇒ at most One 0

Reverse ε-machine: Countably infinite!



Directional Computational Mechanics

- Theorem:

$$\mathbf{E} = I[\mathcal{S}^+; \mathcal{S}^-]$$

- Effective transmission capacity of channel between forward and reverse processes.
- Time agnostic representation: The **BiMachine**.

J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney, "Time's Barbed Arrow: Irreversibility, Crypticity, and Stored Information", Physical Review Letters **103**:9 (2009) 094101.

Information Accessibility

- How hidden is a hidden Process?
- **Crypticity**:

$$\underbrace{\chi = C_\mu}_{\text{Stored Information}} - \underbrace{\mathbf{E}}_{\text{Apparent Information}}$$

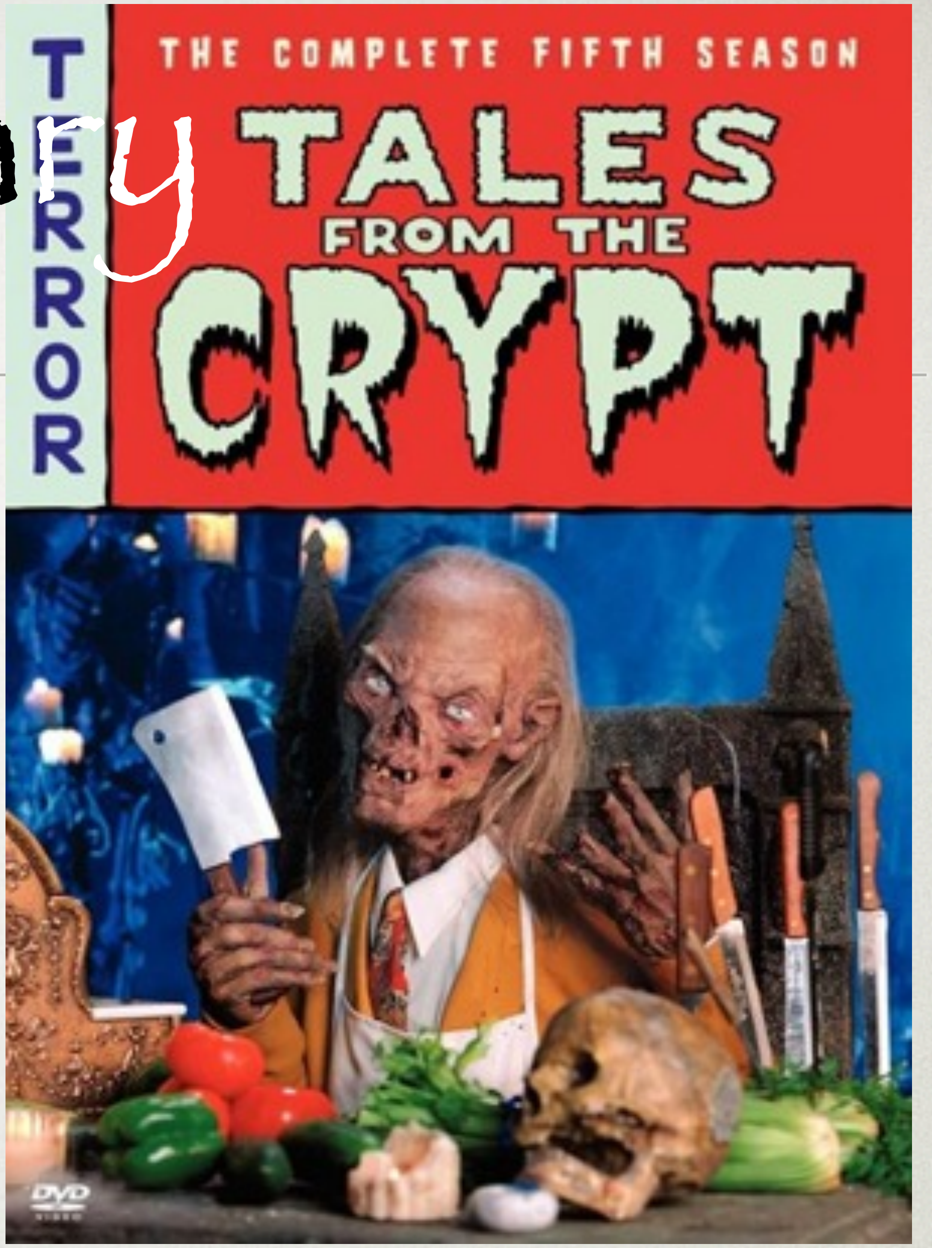
Summary

Information stored in the present

is not

that shared between the past and the future.

Cautionary



ic

- Cryptic Processes: Excess entropy can be arbitrarily small ($\mathbf{E} \approx 0$).
- Even for very structured ($C_\mu \gg 1$) processes.
- **Care** when applying informational analyses to complex systems; esp. mutual information.
- Best to focus on causal architecture, then calculate what you need.

So it goes.

...
We went to the New York World's Fair, saw what the past had been like, according to Ford Motor Car Company and Walt Disney, saw what the future would be like, according to General Motors.

And I asked myself about the present: how wide it was, how deep it was, how much was mine to keep.

Kurt Vonnegut (1922–2007)
Slaughterhouse-Five (1968) p. 23.

Thanks!

http://csc.ucdavis.edu/~chaos/

- J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney, "Time's Barbed Arrow: Irreversibility, Crypticity, and Stored Information", Physical Review Letters **103**:9 (2009) 094101.
- J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield, "Information Accessibility and Cryptic Processes", Journal of Physics A: Math. Theo. **42** (2009) 362002.
- C. J. Ellison, J. R. Mahoney, and J. P. Crutchfield, "Prediction, Retrodiction, and the Amount of Information Stored in the Present", Journal of Statistical Physics **137**:6 (2009) 1005-1034.
- J. Mahoney, C. J. Ellison, and J. P. Crutchfield, "Information Accessibility and Cryptic Processes: Linear Combinations of Causal States", arxiv.org:0906.5099 [cond-mat].
- J. P. Crutchfield, C. J. Ellison, J. R. Mahoney, and R. G. James, "Synchronization and Control in Intrinsic and Designed Computation: An Information-Theoretic Analysis of Competing Models of Stochastic Computation", CHAOS **20**:3 (2010) 037105.

Extras

Time's Barbed Arrow: The Past & the Future in the Present

Jim Crutchfield
Complexity Sciences Center
Physics Department
University of California at Davis

Theoretical Neuroscience Seminar
Center for Neuroscience
University of California at Davis

18 February 2011

Joint work with Chris Ellison (UC Davis Physics) & John Mahoney (UC Merced)

The Past and the Future in the Present:

Directional Computational Mechanics

Jim Crutchfield
Complexity Sciences Center
Physics Department
University of California at Davis

Workshop on
Modeling Dynamical Systems
University of California at Davis

13 march 2010

Joint work with Chris Ellison (UCD Physics) & John Mahoney (UCD Physics)

Time's Barbed Arrow: The Past & the Future in the Present

Jim Crutchfield
Complexity Sciences Center
Physics Department
University of California at Davis

Redwood Center for Theoretical Neuroscience
University of California at Berkeley

8 October 2010

Joint work with Chris Ellison (UCD Physics) & John Mahoney (UCD Physics)

The Past & the Future in the Present

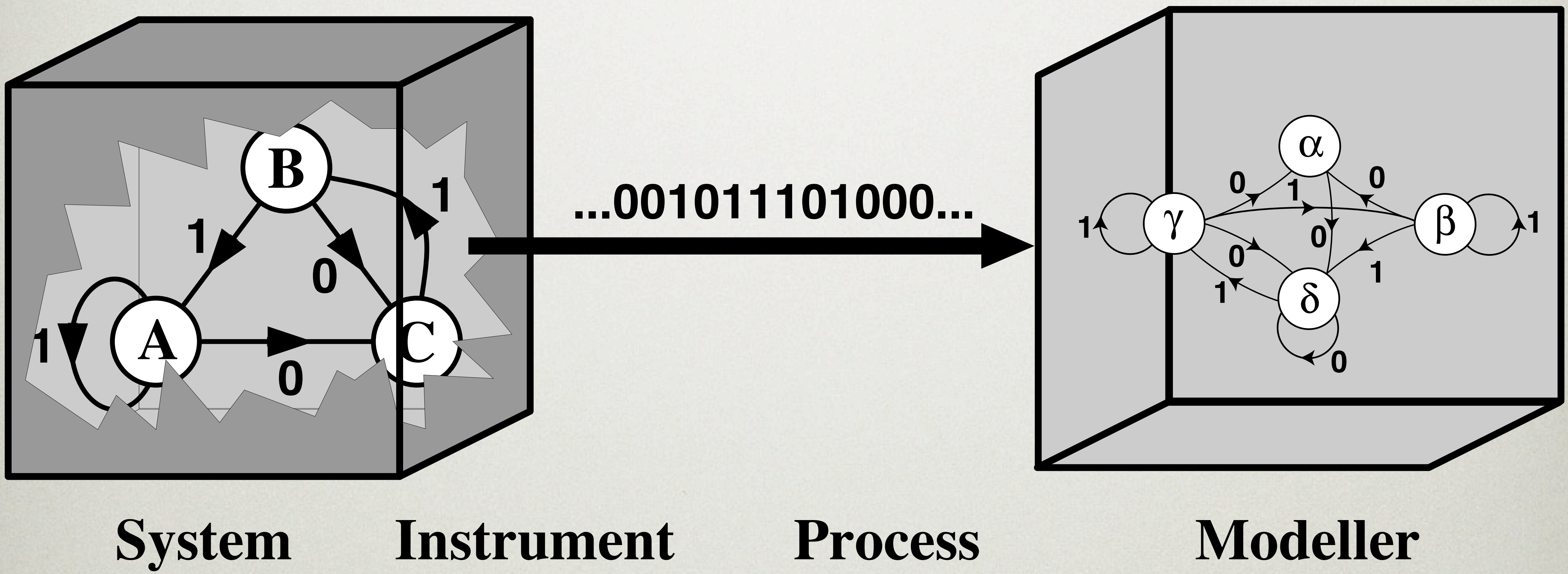
Jim Crutchfield
Complexity Sciences Center
Physics Department
University of California at Davis

Colloquium
Santa Fe Institute
Santa Fe, New Mexico

10 March 2011

Joint work with Chris Ellison (UC Davis) & John Mahoney (UC Merced)

The Learning Channel



N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, "Geometry from a Time Series", Physical Review Letters **45** (1980) 712.
J. P. Crutchfield, B. S. McNamara, "Equations of Motion from a Data Series", Complex Systems **1** (1987) 417-452.

Computational Mechanics

- Theorem (Causal Shielding):

$$\Pr(\overleftarrow{X}, \overrightarrow{X} | \mathcal{S}) = \Pr(\overleftarrow{X} | \mathcal{S}) \Pr(\overrightarrow{X} | \mathcal{S})$$

- Theorem (Optimal Prediction):

$$\Pr(\overrightarrow{X} | \mathcal{S}) = \Pr(\overrightarrow{X} | \overleftarrow{X})$$

- Corollary (Capture All Shared Information):

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

(Prescient models)

- Theorem: ε-Machine is smallest prescient model

$$C_\mu \equiv H[\mathcal{S}] \leq H[\widehat{\mathcal{R}}]$$

Computational Mechanics

- A prediction: Map from a past to possible futures

$$\Pr(\overrightarrow{X}|\overleftarrow{x})$$

- A good predictor $\widehat{\mathcal{R}}$ captures all of the *predictable information* between past and future:

$$\mathbf{E} = I[\widehat{\mathcal{R}};\overrightarrow{X}]$$

- Modeling:
  - Make good predictions, but also
  - Represent underlying mechanisms

J. P. Crutchfield, K. Young, "Inferring Statistical Complexity", Physical Review Letters **63** (1989) 105-108.

Focus Problem: $\mathbf{E}$ Versus C_μ

- Can get h_μ and C_μ directly from ε-Machine.
- How to calculate $\mathbf{E}$ from ε-Machine?

- Return to the larger issues at the beginning (relating modeling and prediction), but with a new “invariant”: information accessibility.

Focus Problem: $\mathbf{E}$ Versus C_μ

- Known:
  - Range-R spin systems:

$$C_\mu = \mathbf{E} + R h_\mu$$

J. P. Crutchfield and D. P. Feldman,
"Statistical Complexity of Simple One-Dimensional Spin Systems",
Physical Review E **55**:2 (1997) R1239-R1243.

- Theorem: $\mathbf{E} \leq C_\mu$

C. R. Shalizi and J. P. Crutchfield, "Computational Mechanics:
Pattern and Prediction, Structure and Simplicity", J. Stat. Phys.
104 (2001) 817-879.

Directional Computational Mechanics

- Temporal asymmetry:

$$C_\mu^- \neq C_\mu^+$$

- Causal Irreversibility:

$$\begin{aligned}\Xi &\equiv C_\mu^+ - C_\mu^- \\ &= H[\mathcal{S}^+|\mathcal{S}^-] - H[\mathcal{S}^-|\mathcal{S}^+]\end{aligned}$$

- Time-symmetric component ($\mathbf{E}$) cancels!

J. P. Crutchfield, "Semantics and Thermodynamics", in **Nonlinear Modeling and Forecasting**, M. Casdagli and S. Eubank, editors, Addison-Wesley, Reading, Massachusetts (1992) 317-359.

Directional Computational Mechanics

- Corollary:

$$C_\mu^\pm = \mathbf{E} + H[\mathcal{S}^+|\mathcal{S}^-] + H[\mathcal{S}^-|\mathcal{S}^+]$$

- Crypticity:

$$\chi \equiv H[\mathcal{S}^+|\mathcal{S}^-] + H[\mathcal{S}^-|\mathcal{S}^+]$$

Distance between measurements & model:

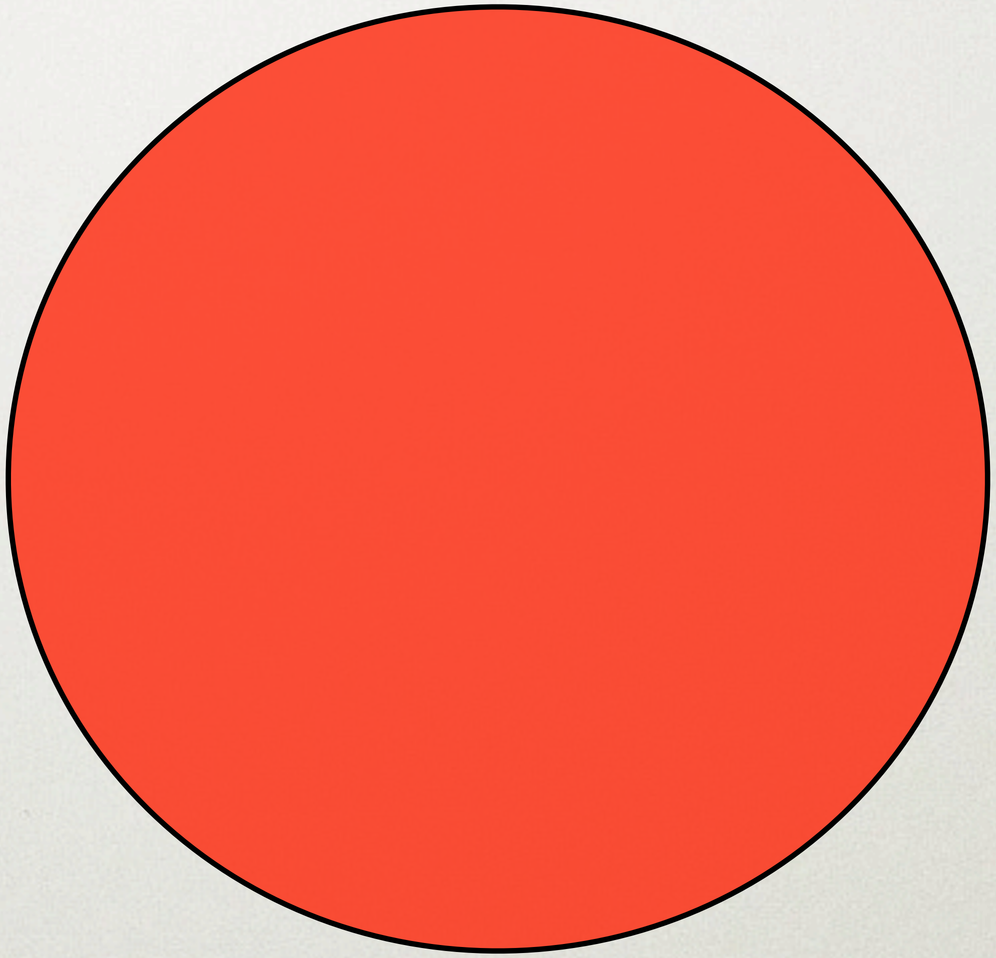
$$d(X,Y) = H[X|Y] + H[Y|X]$$

Degree to which internal information is hidden.

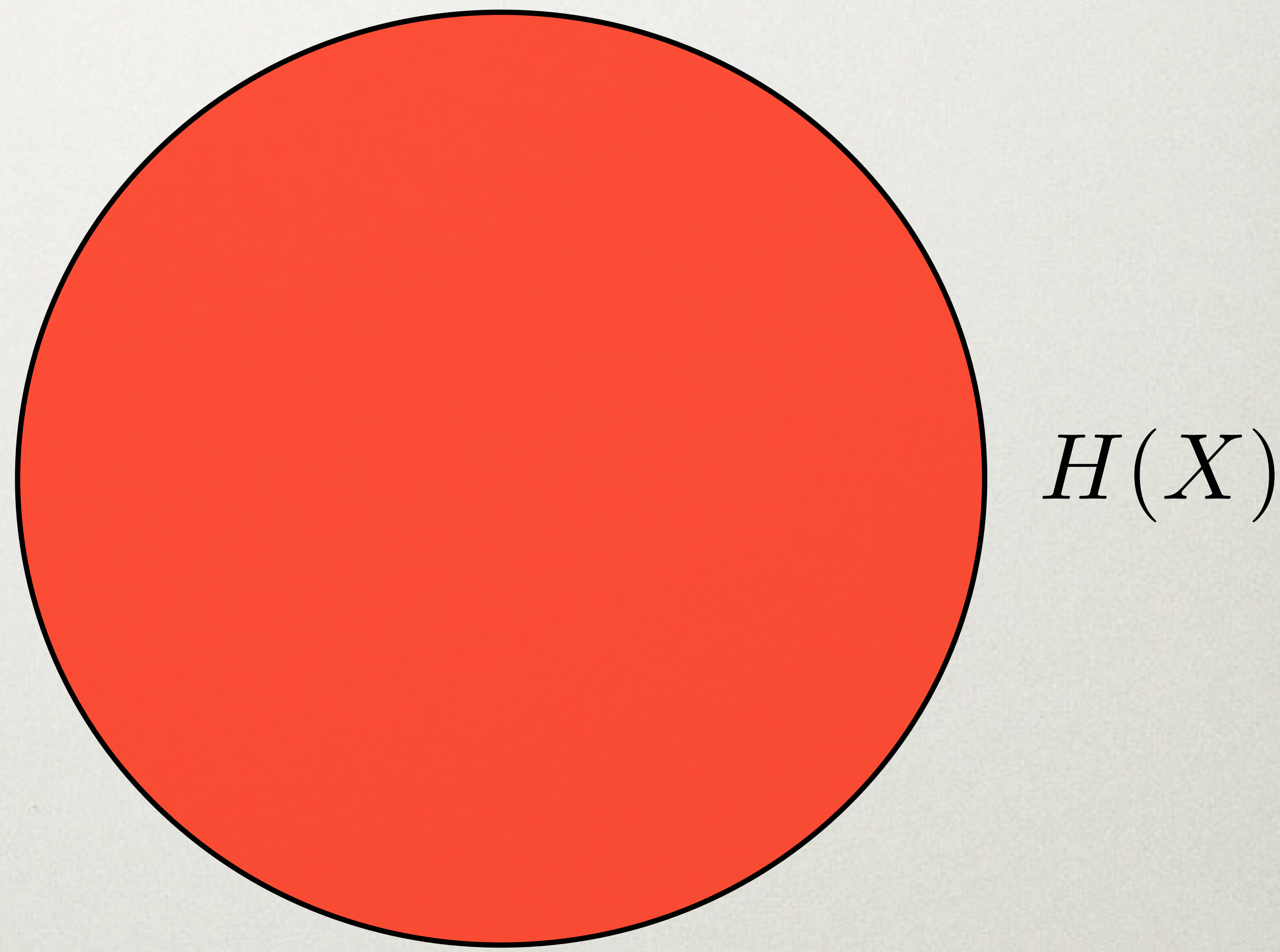
Information inaccessibility!

Information Diagram

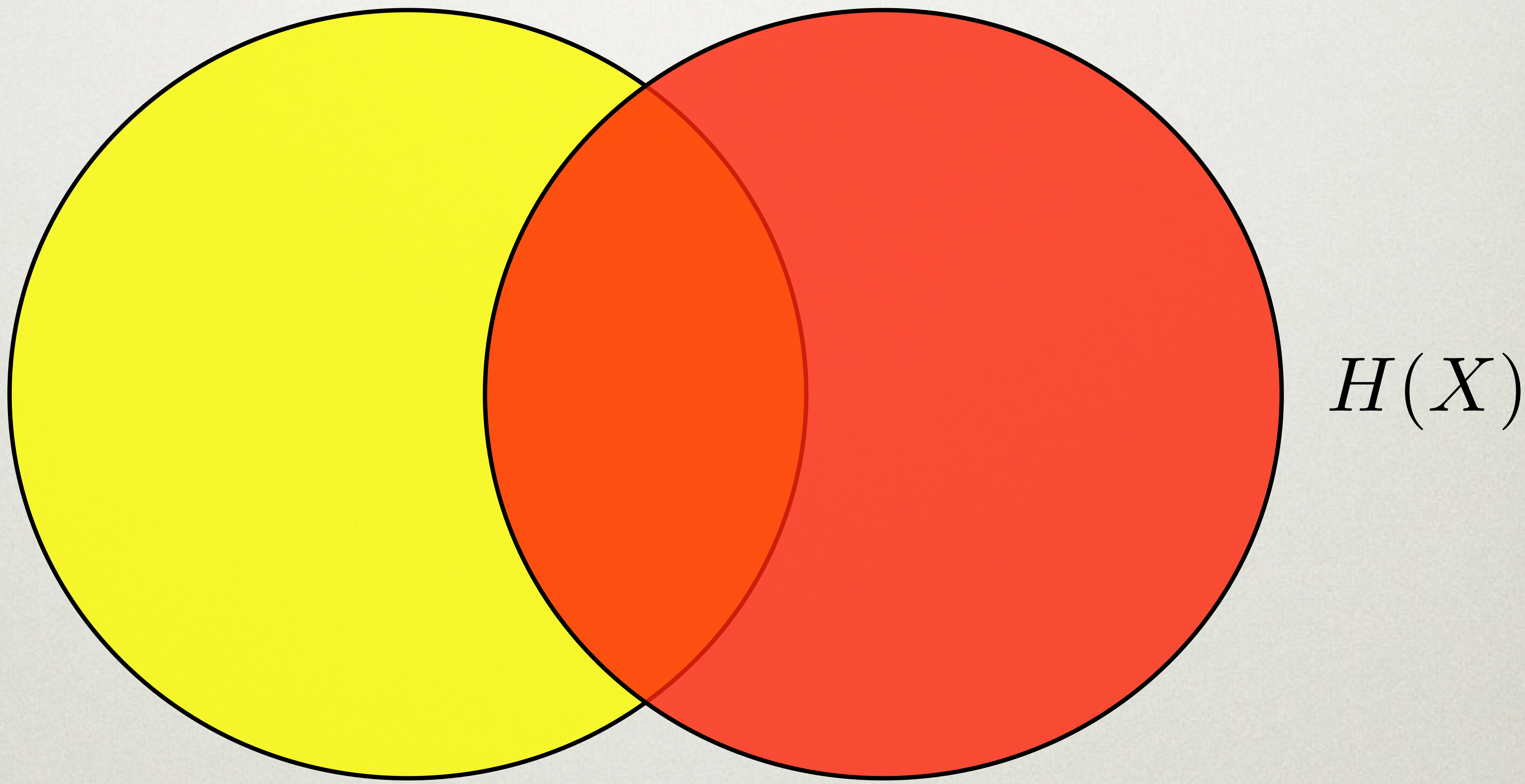
Information Diagram



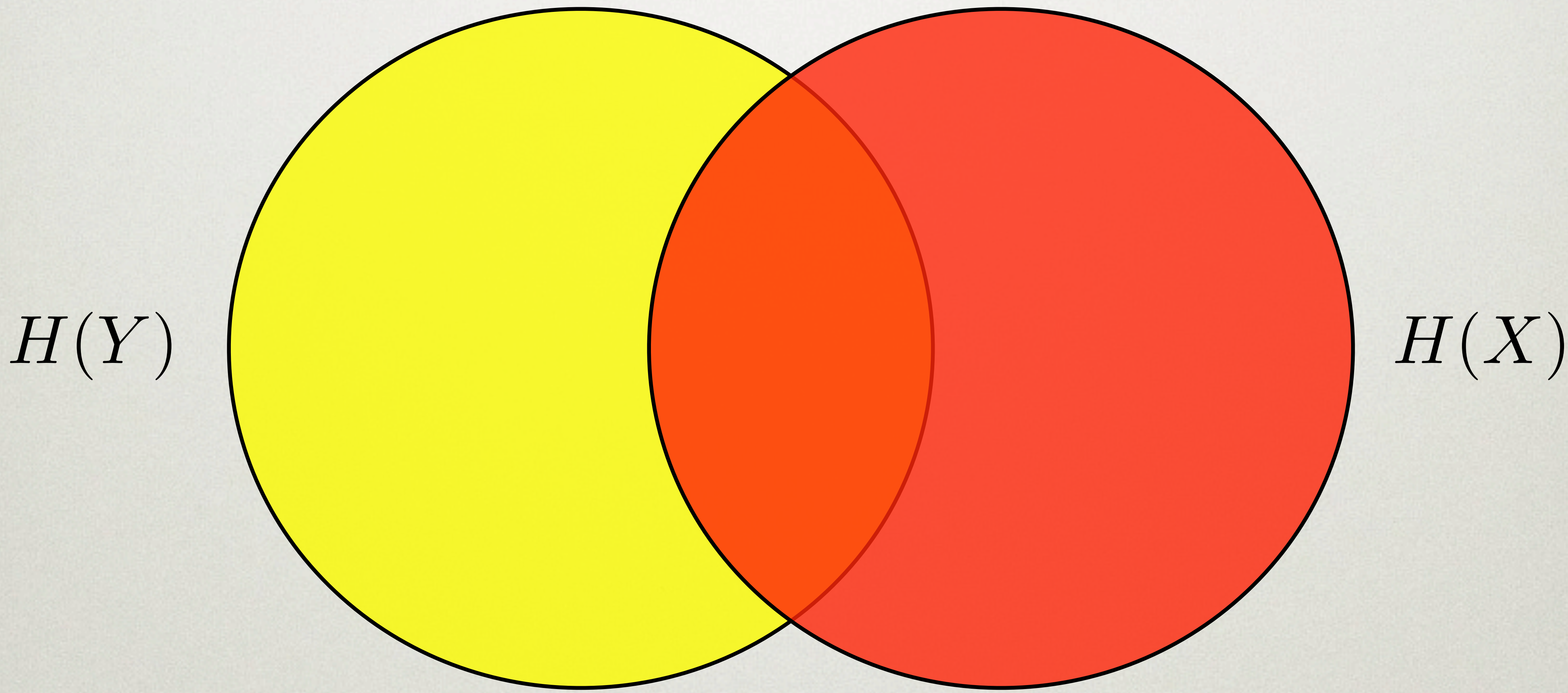
Information Diagram



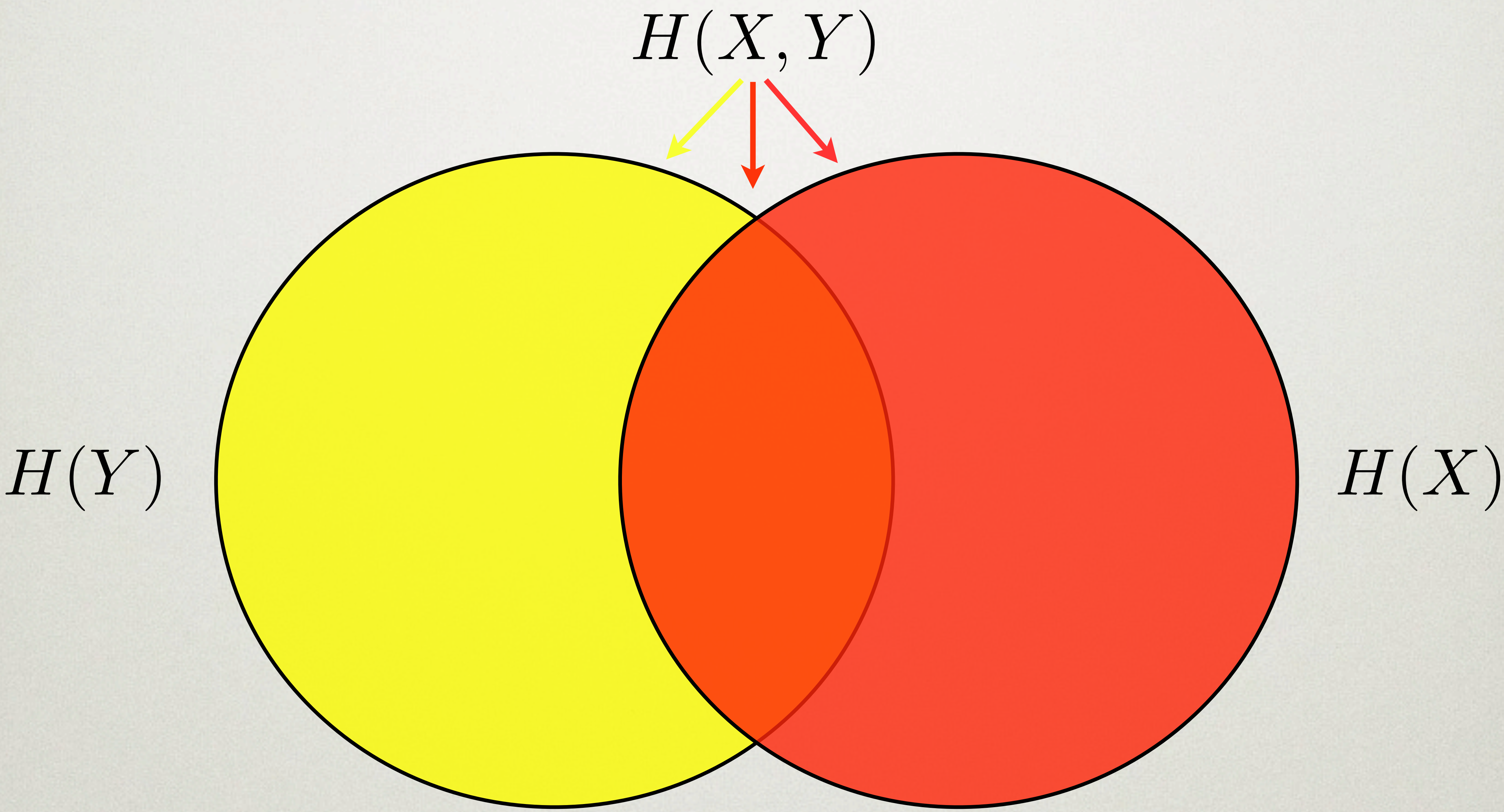
Information Diagram



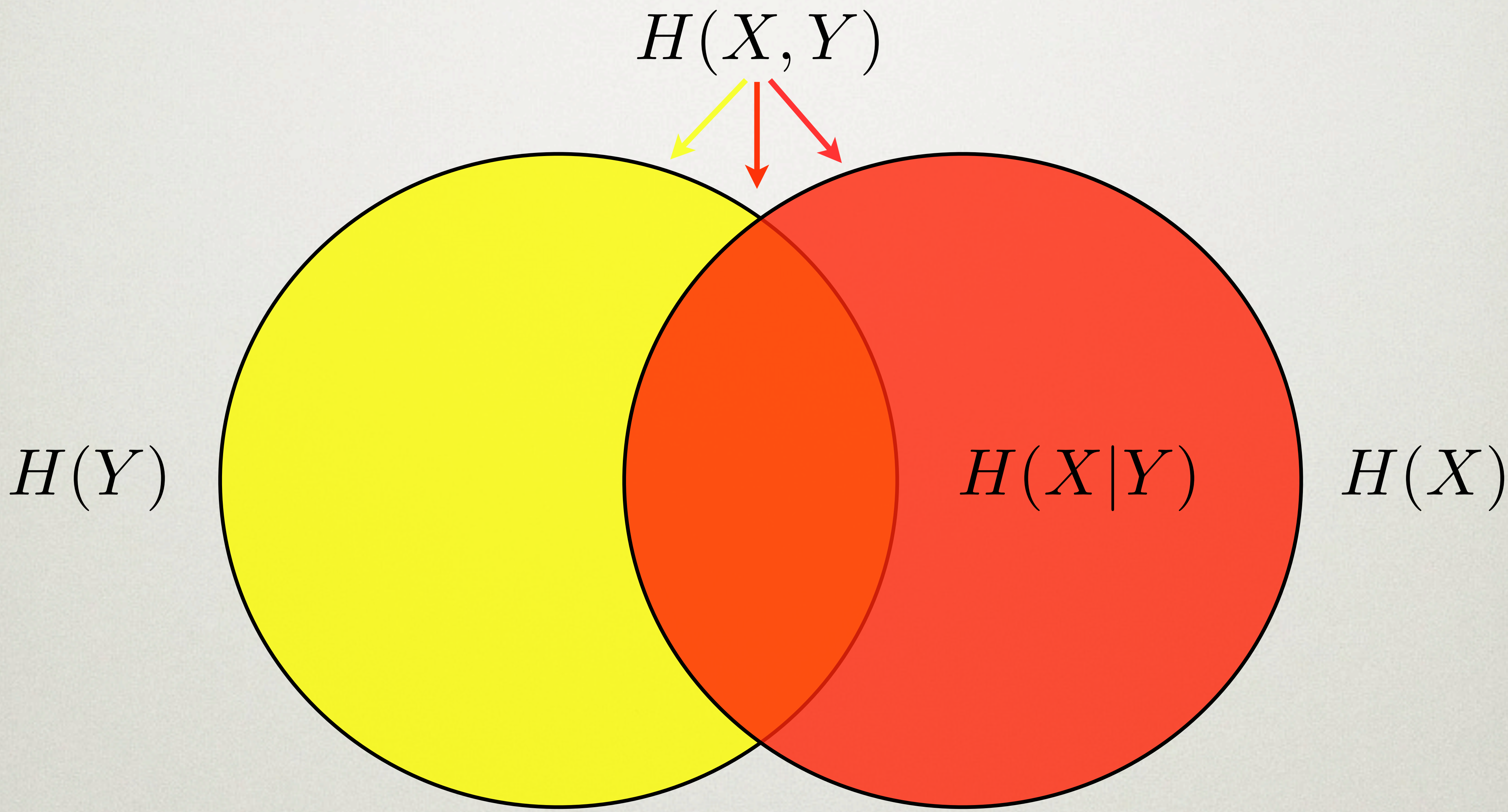
Information Diagram



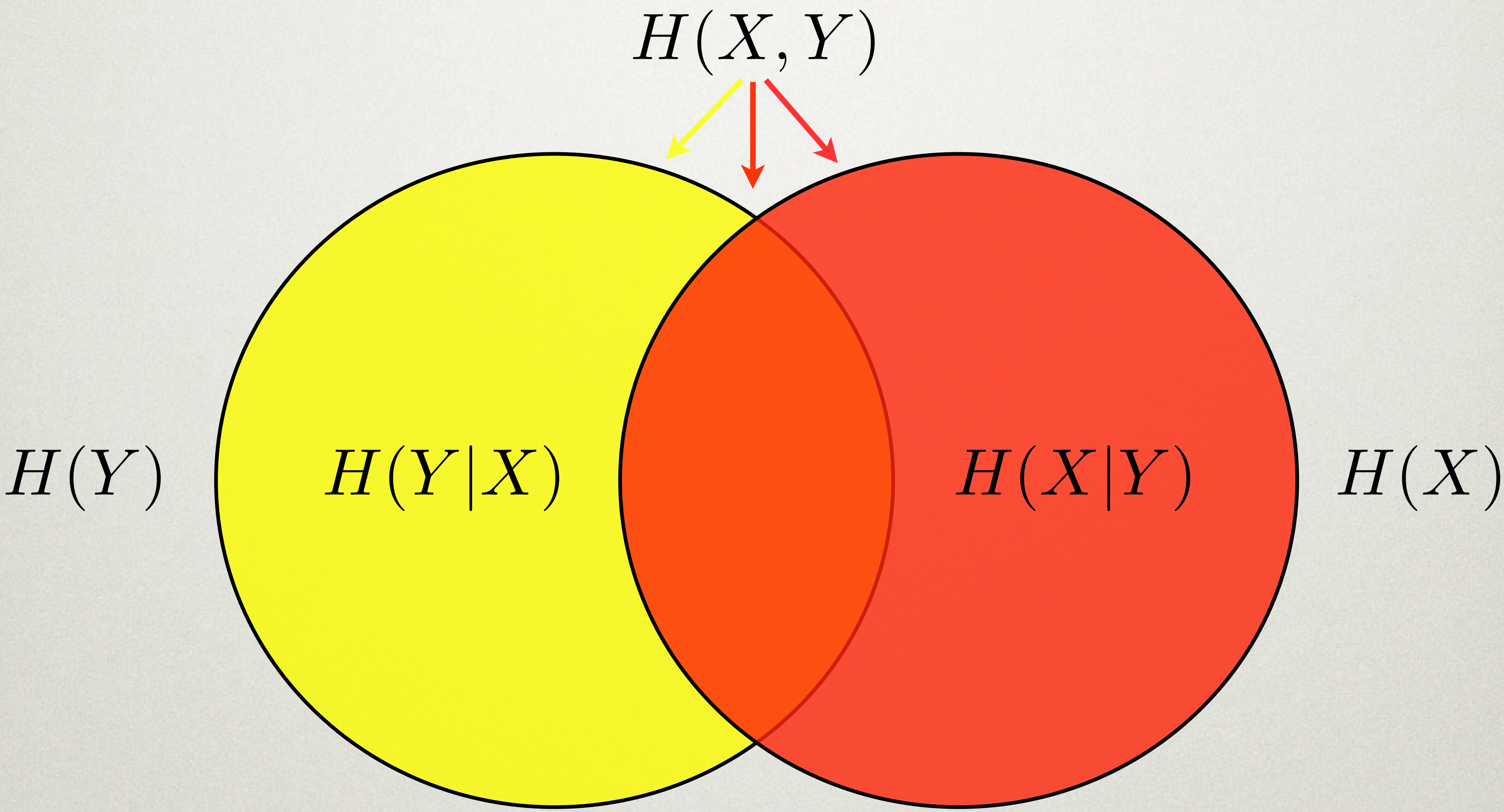
Information Diagram



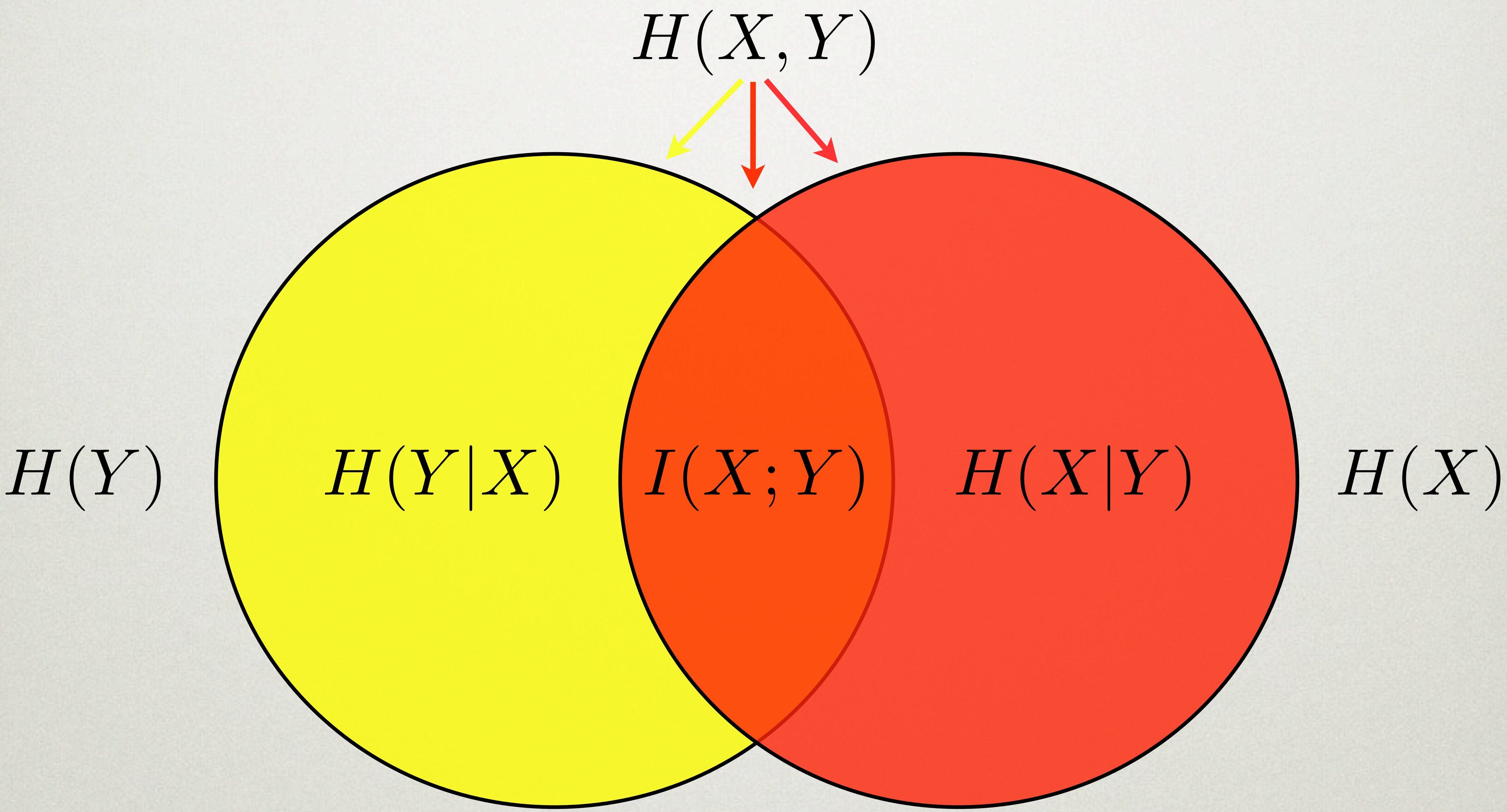
Information Diagram



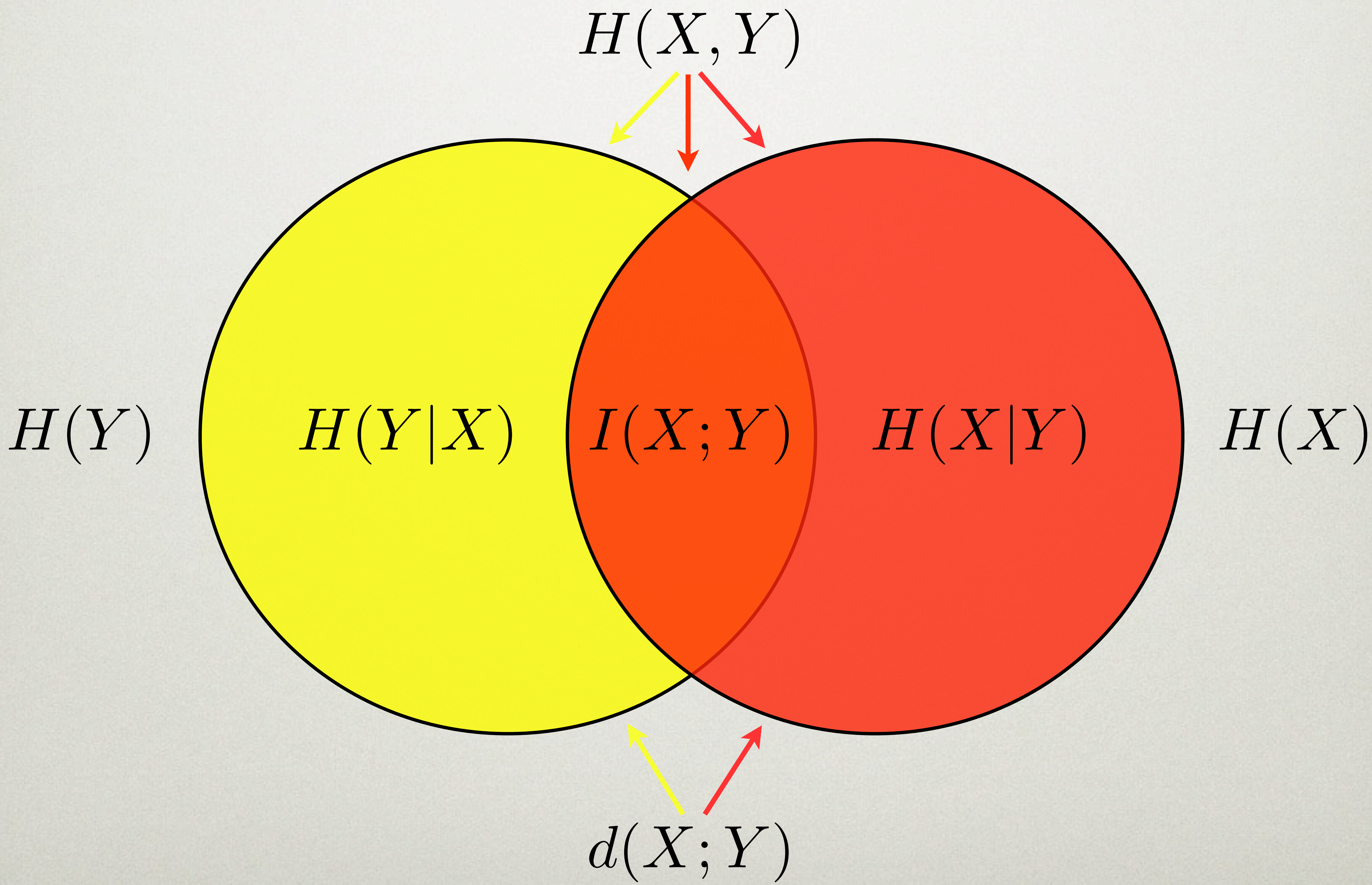
Information Diagram



Information Diagram

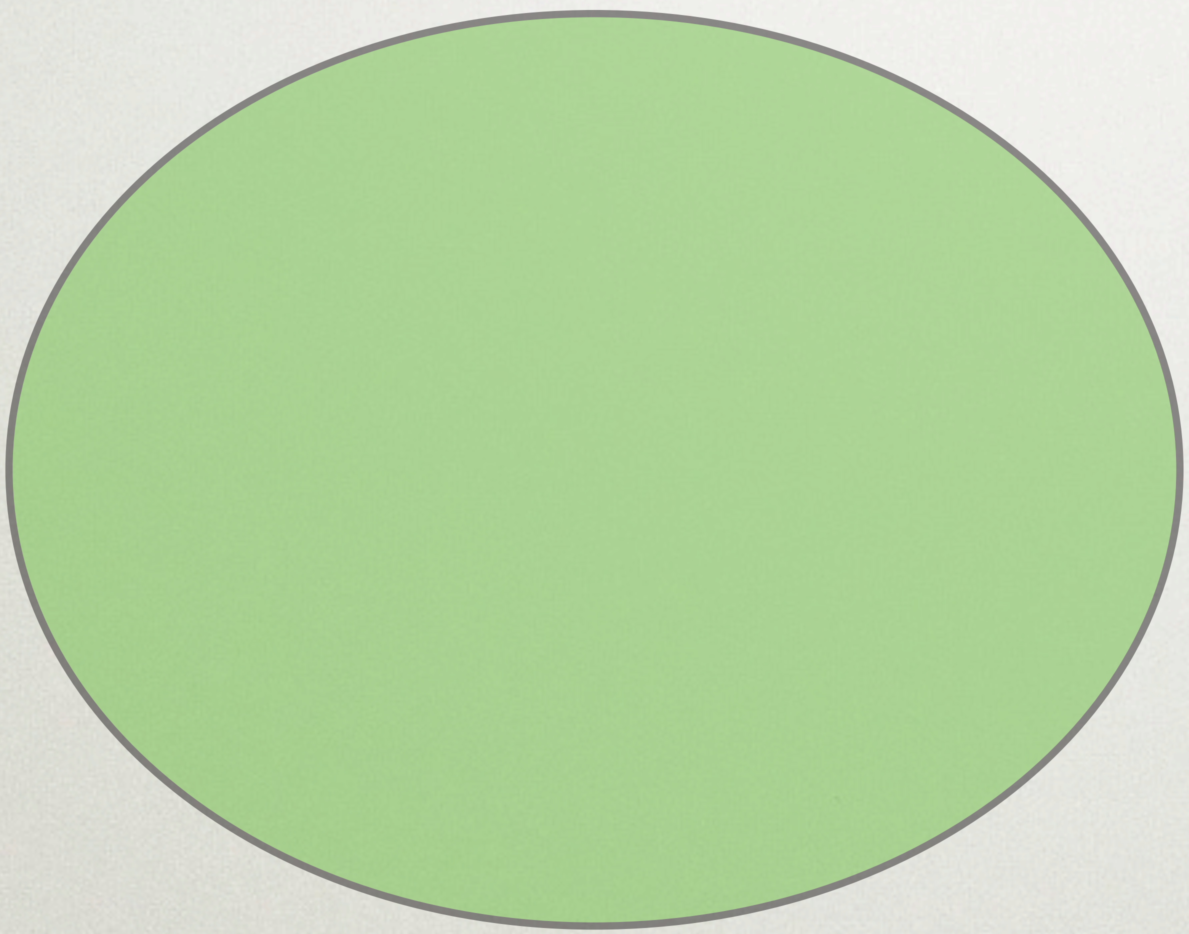


Information Diagram



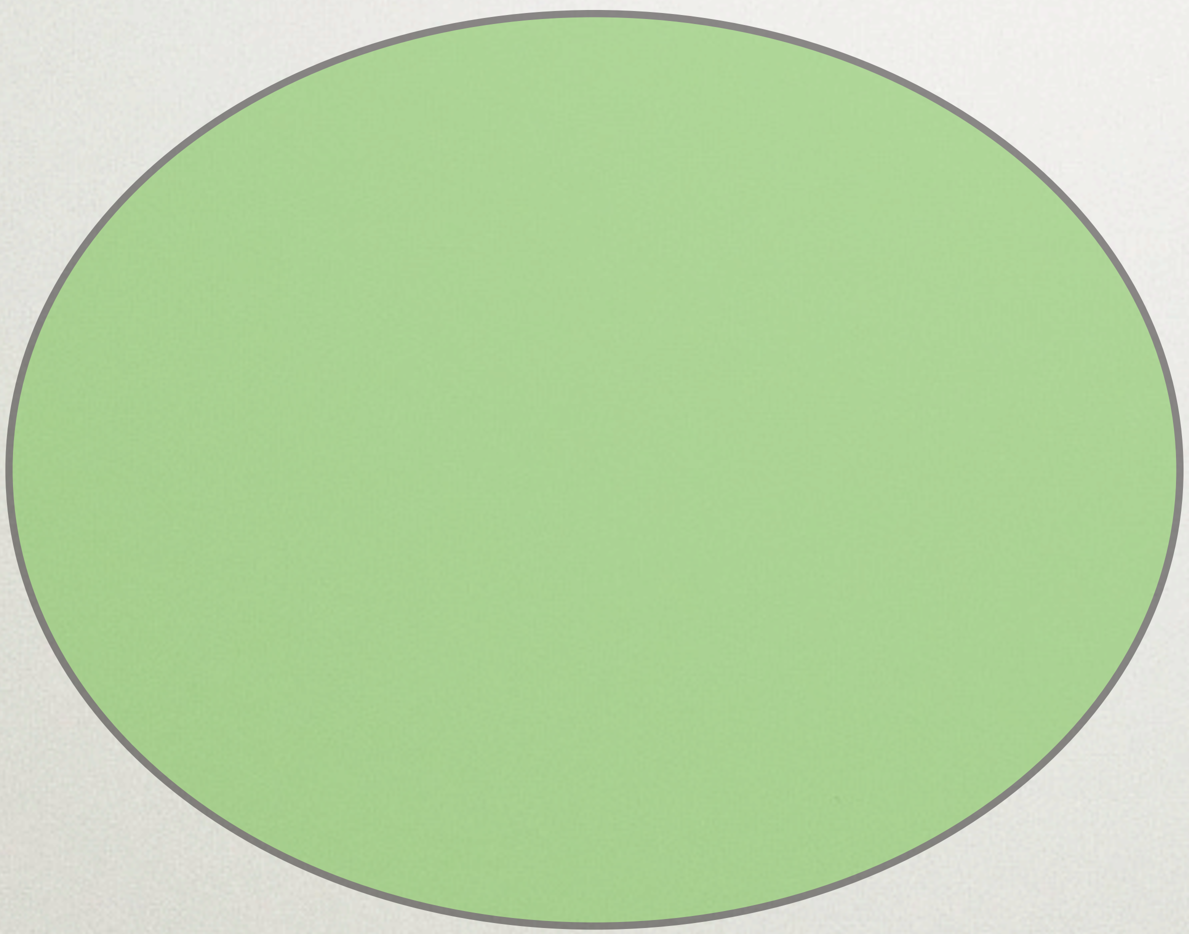
ε-machine information Diagram

ε-machine information Diagram

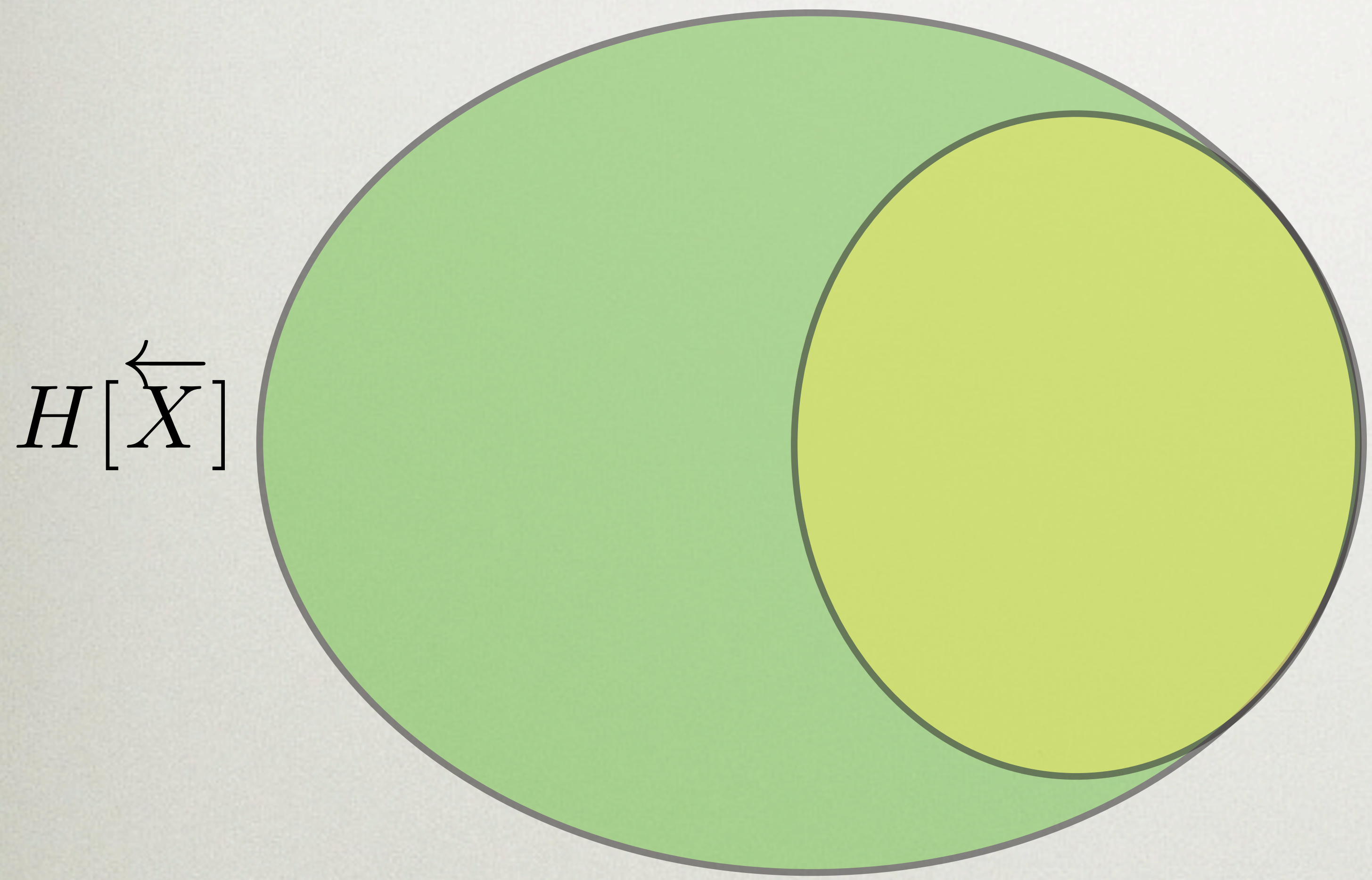


ε-machine information Diagram

$H[\overleftarrow{X}]$



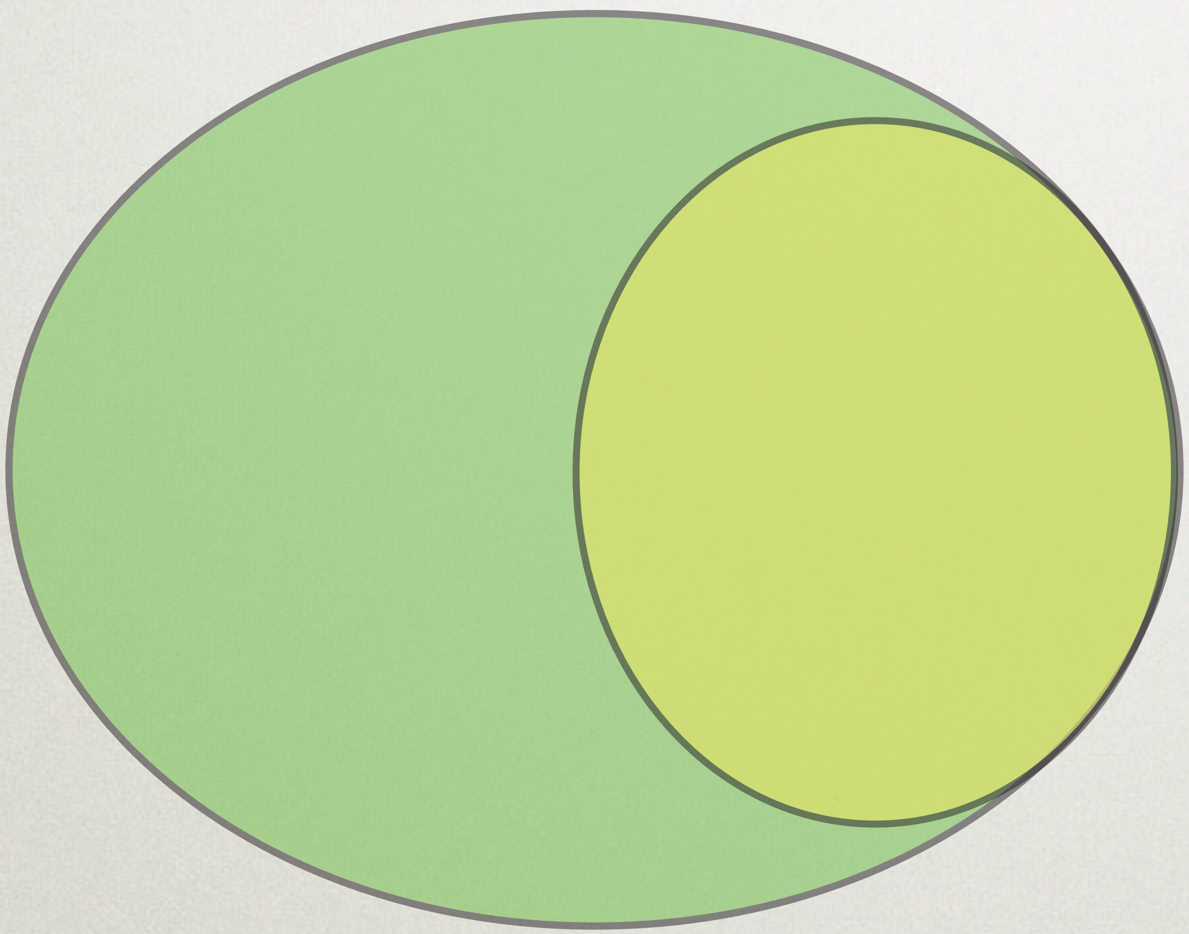
ε-machine information Diagram



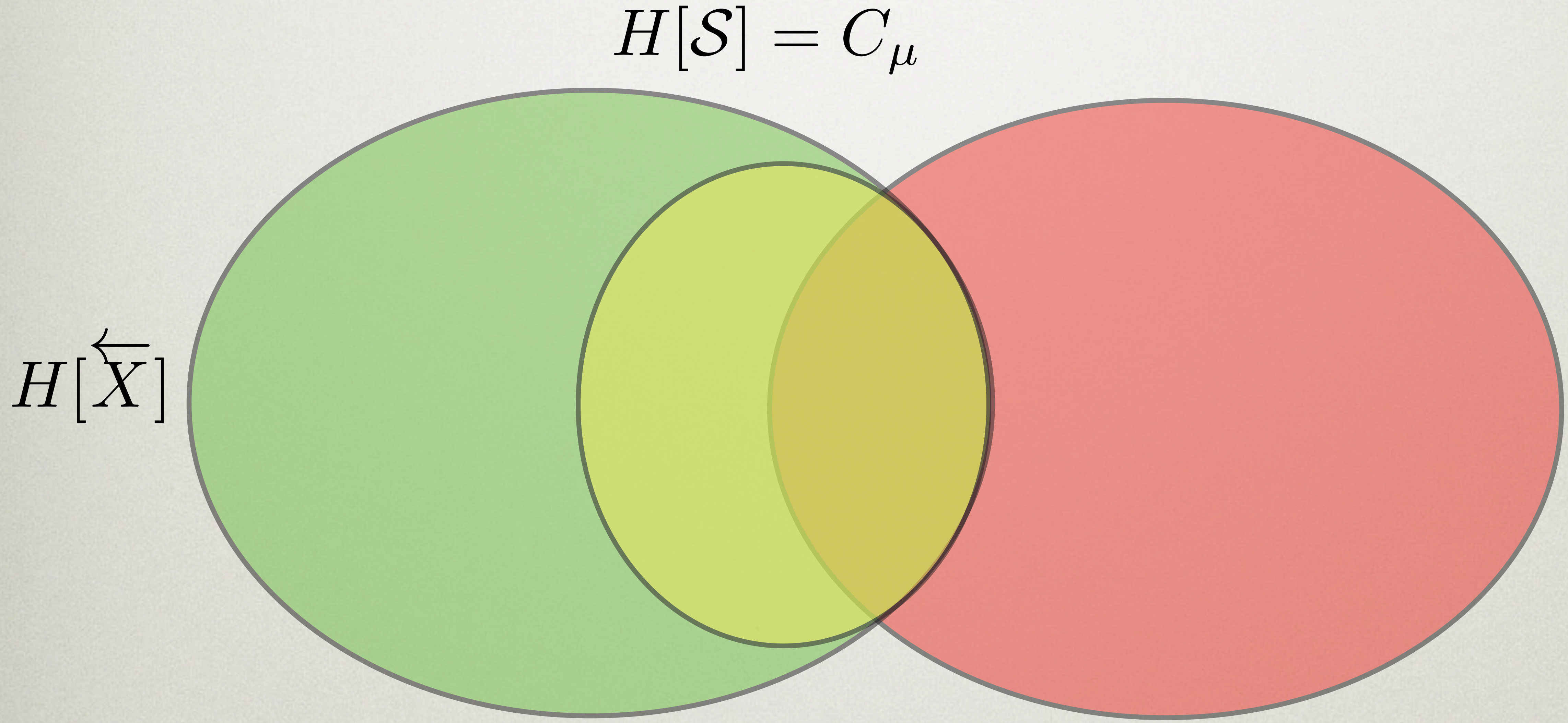
ε-machine information Diagram

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

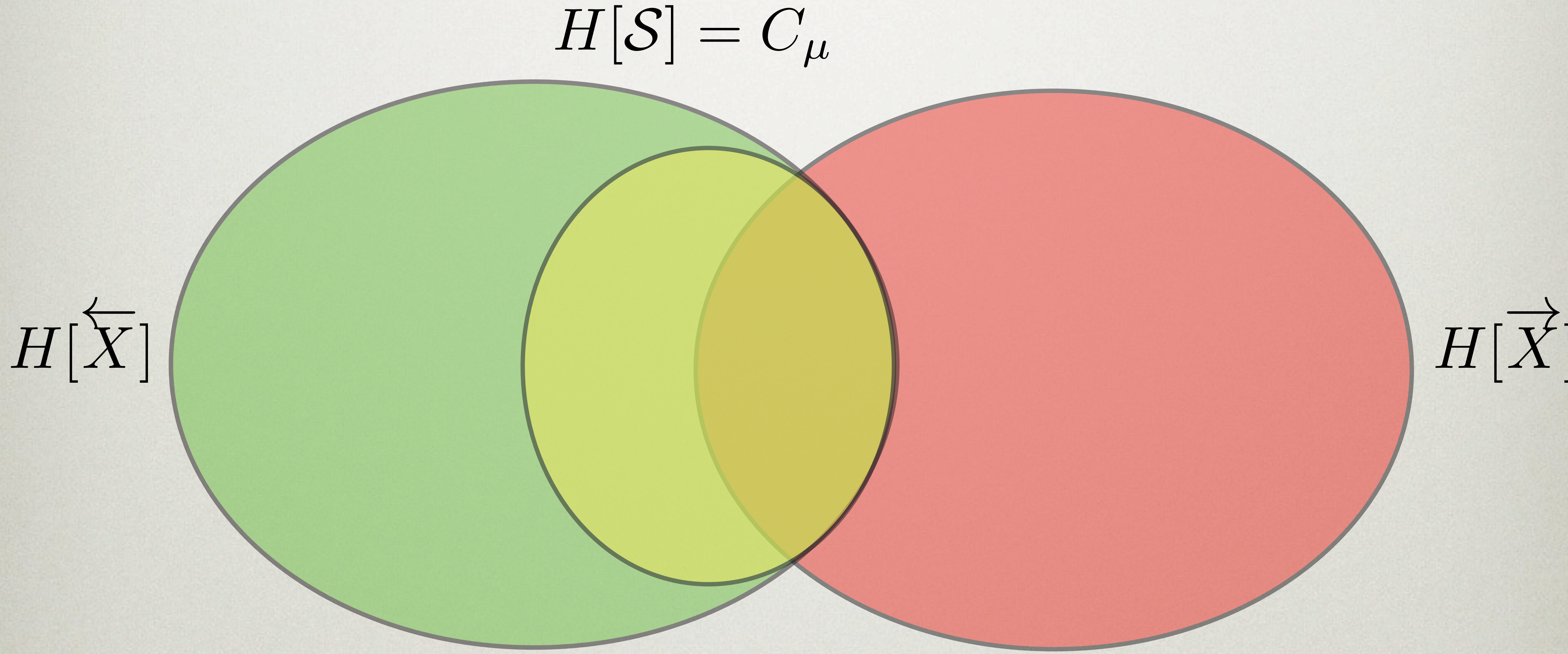
$H[\overleftarrow{X}]$



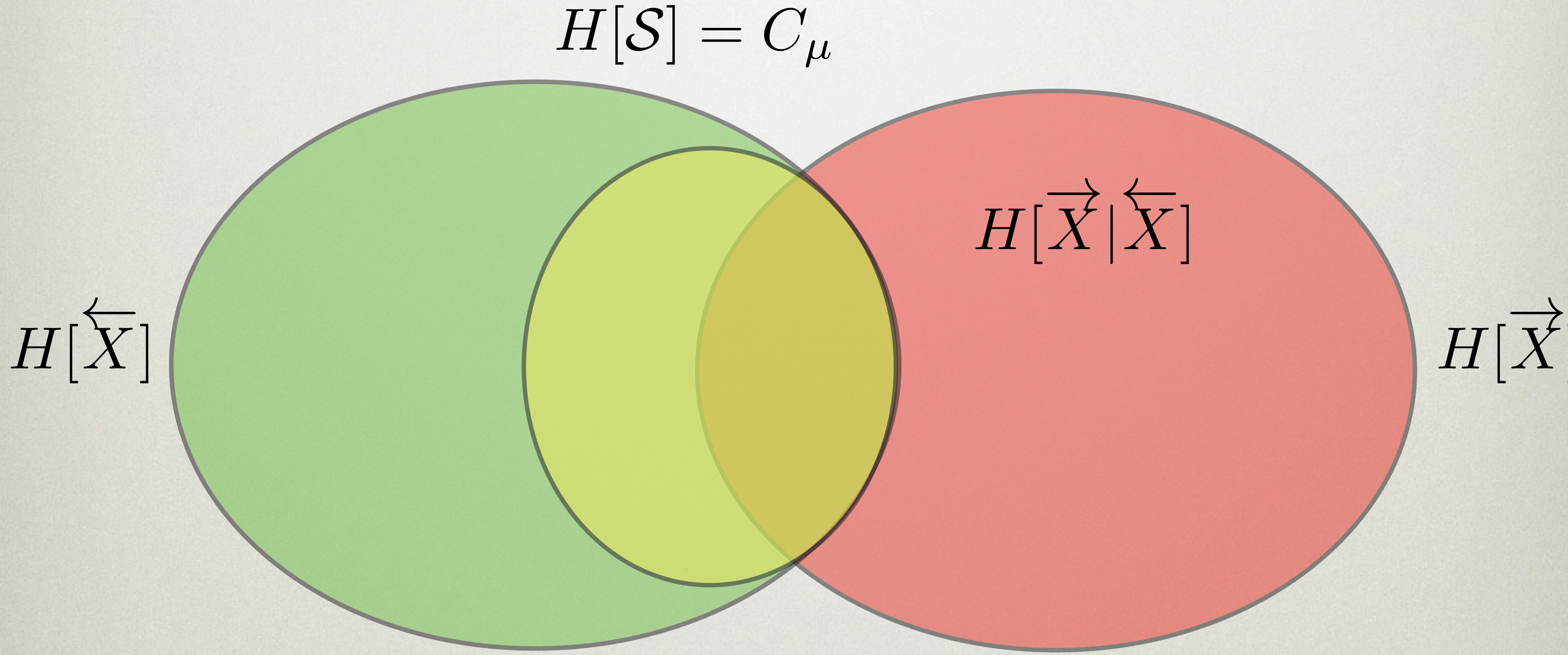
ε-machine information Diagram



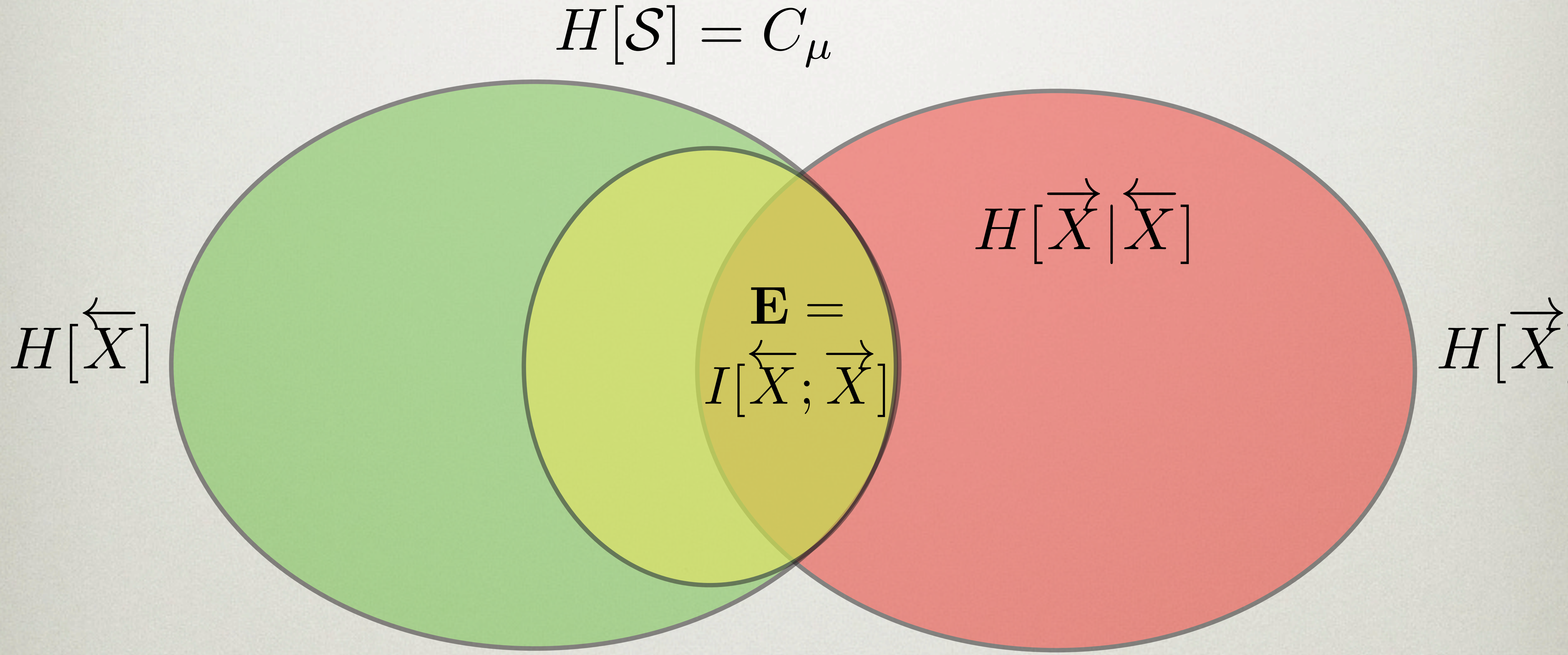
ε-machine information Diagram



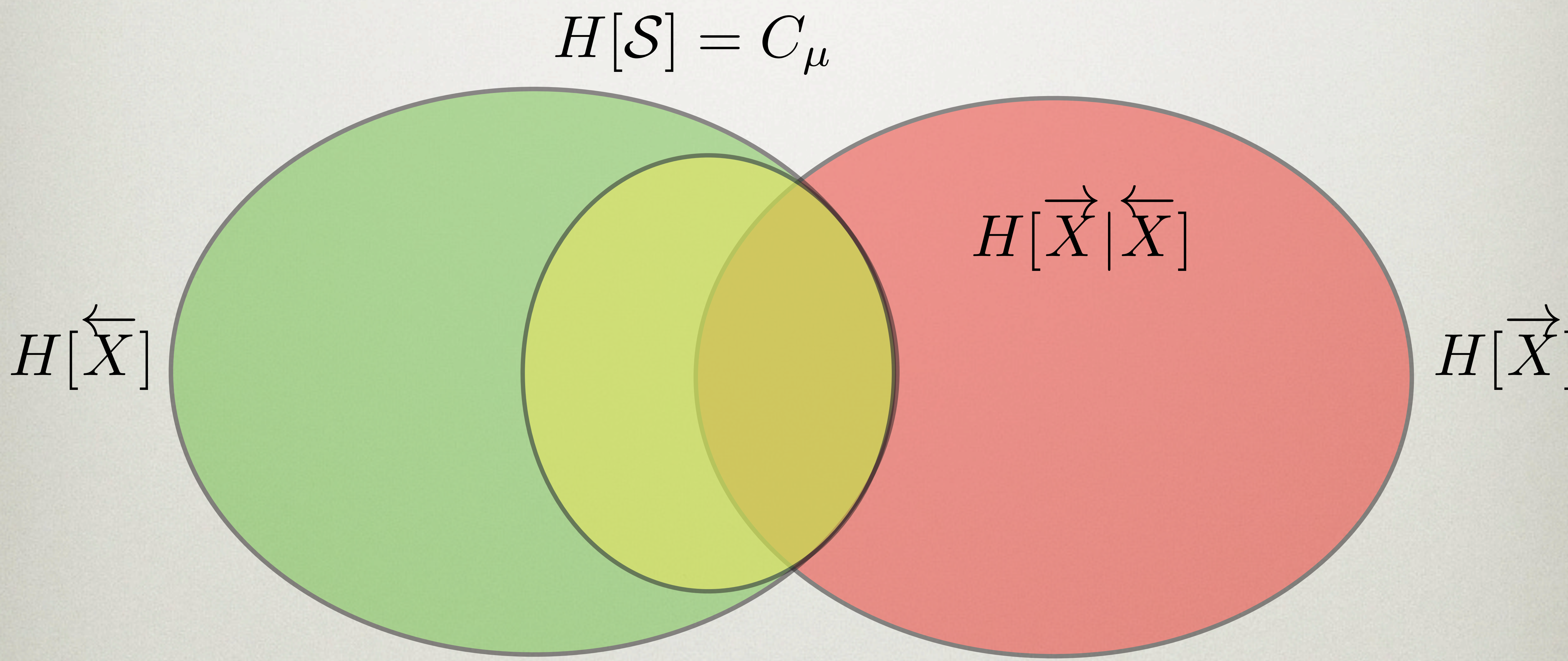
ε-machine information Diagram



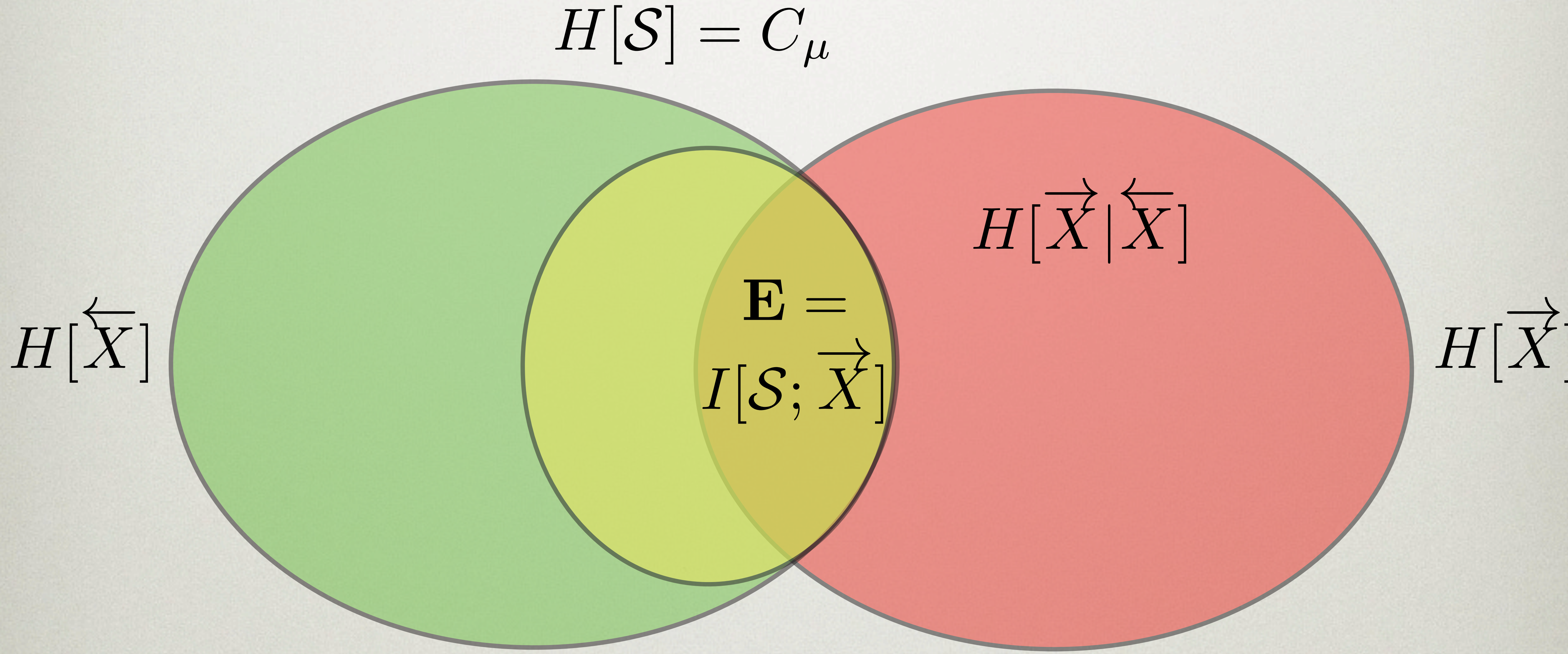
ε-machine information Diagram



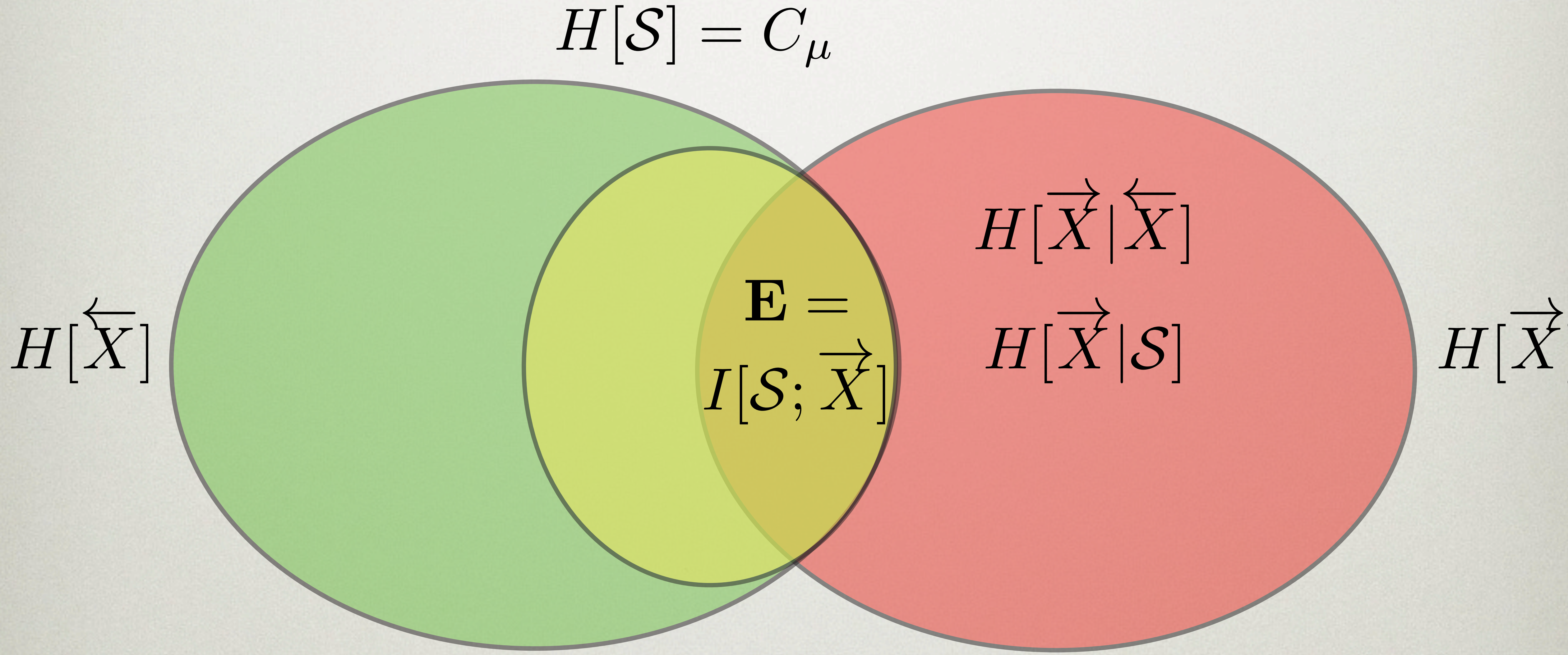
ε-machine information Diagram



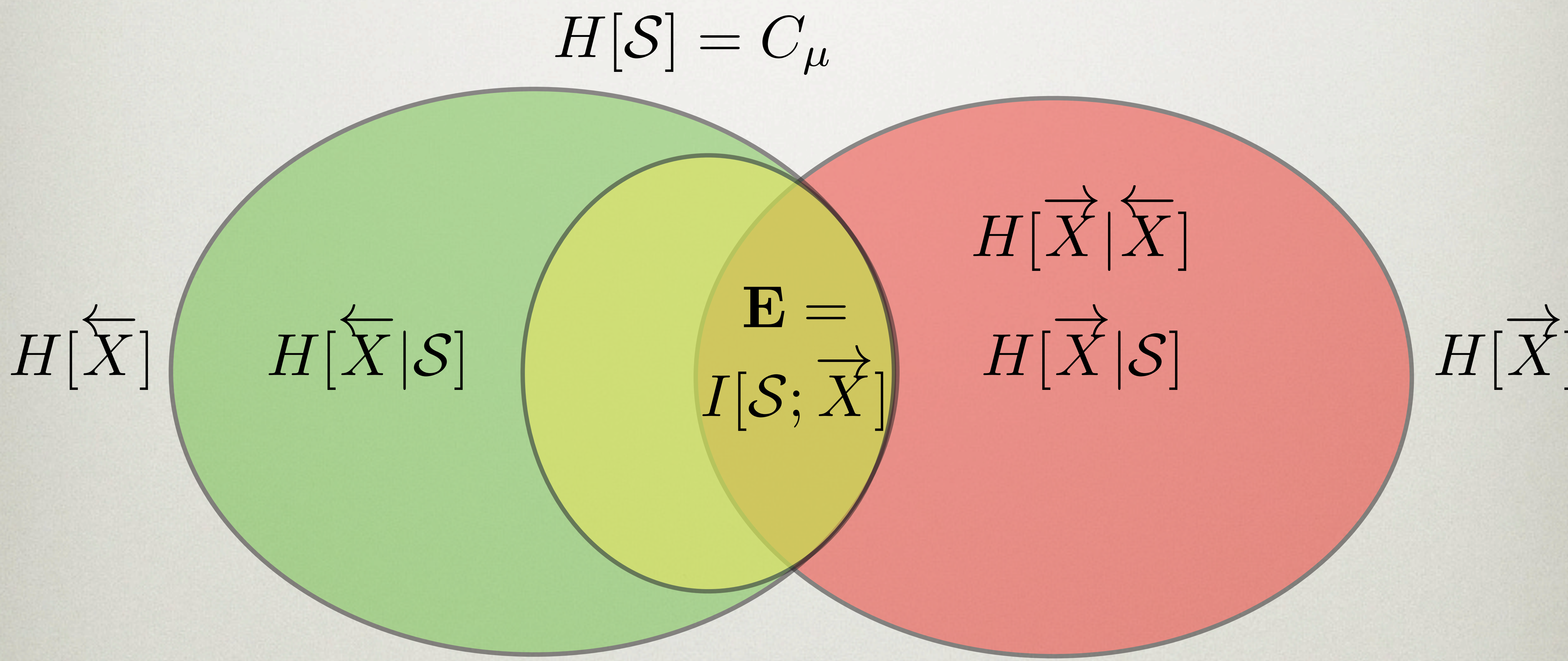
ε-machine information Diagram



ε-machine information Diagram

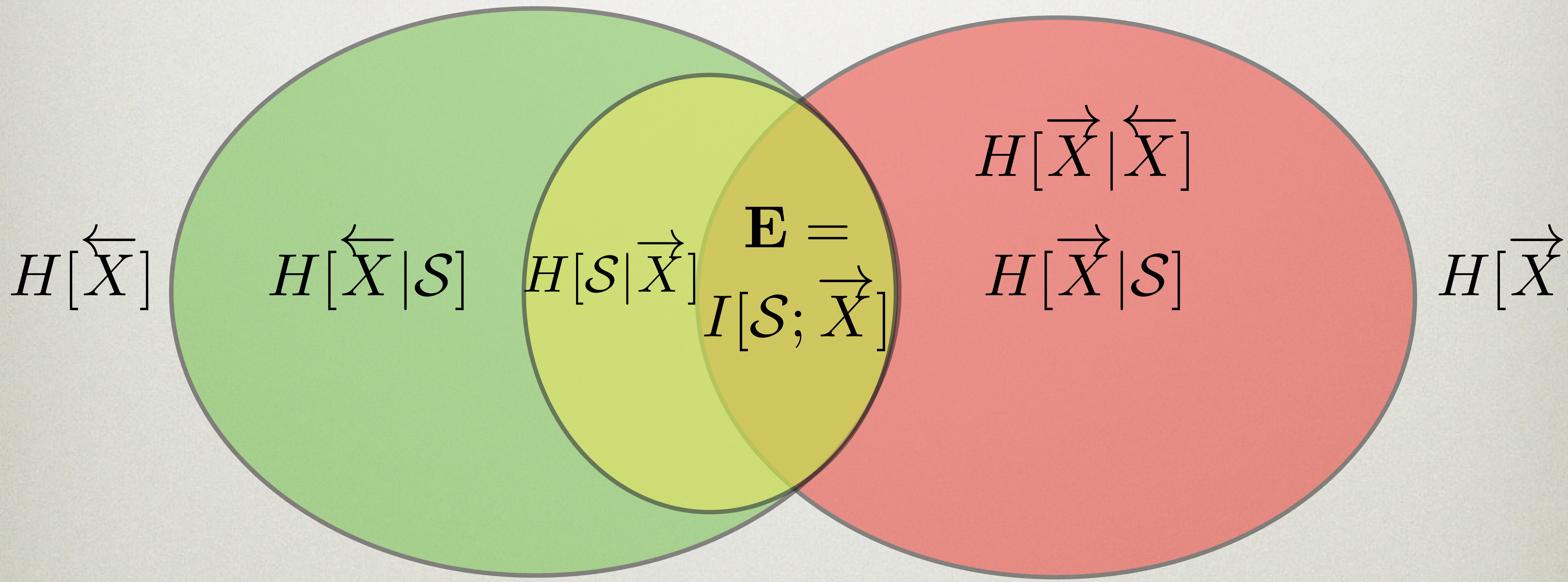


ε-machine information Diagram

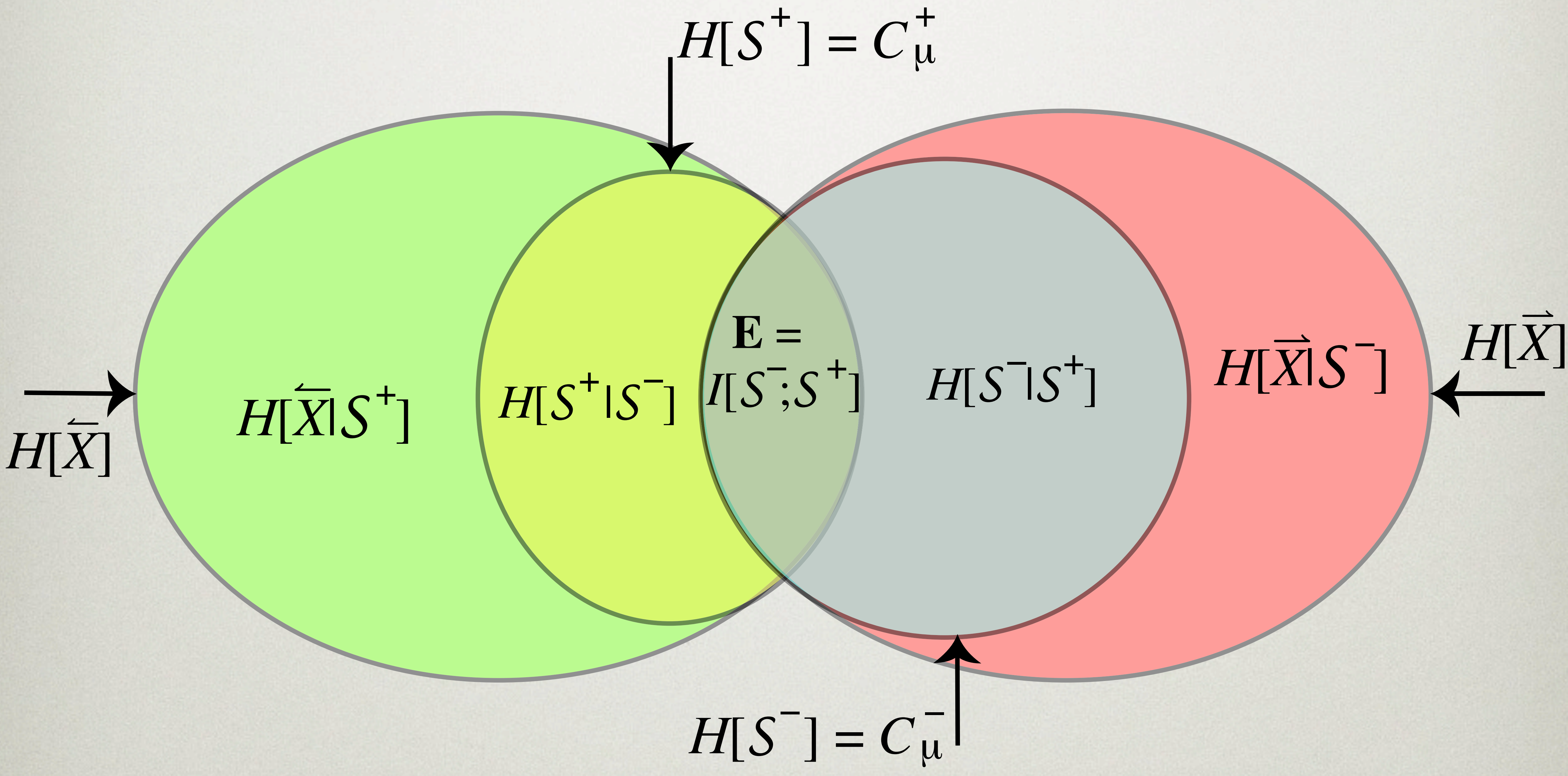


ε-machine information Diagram

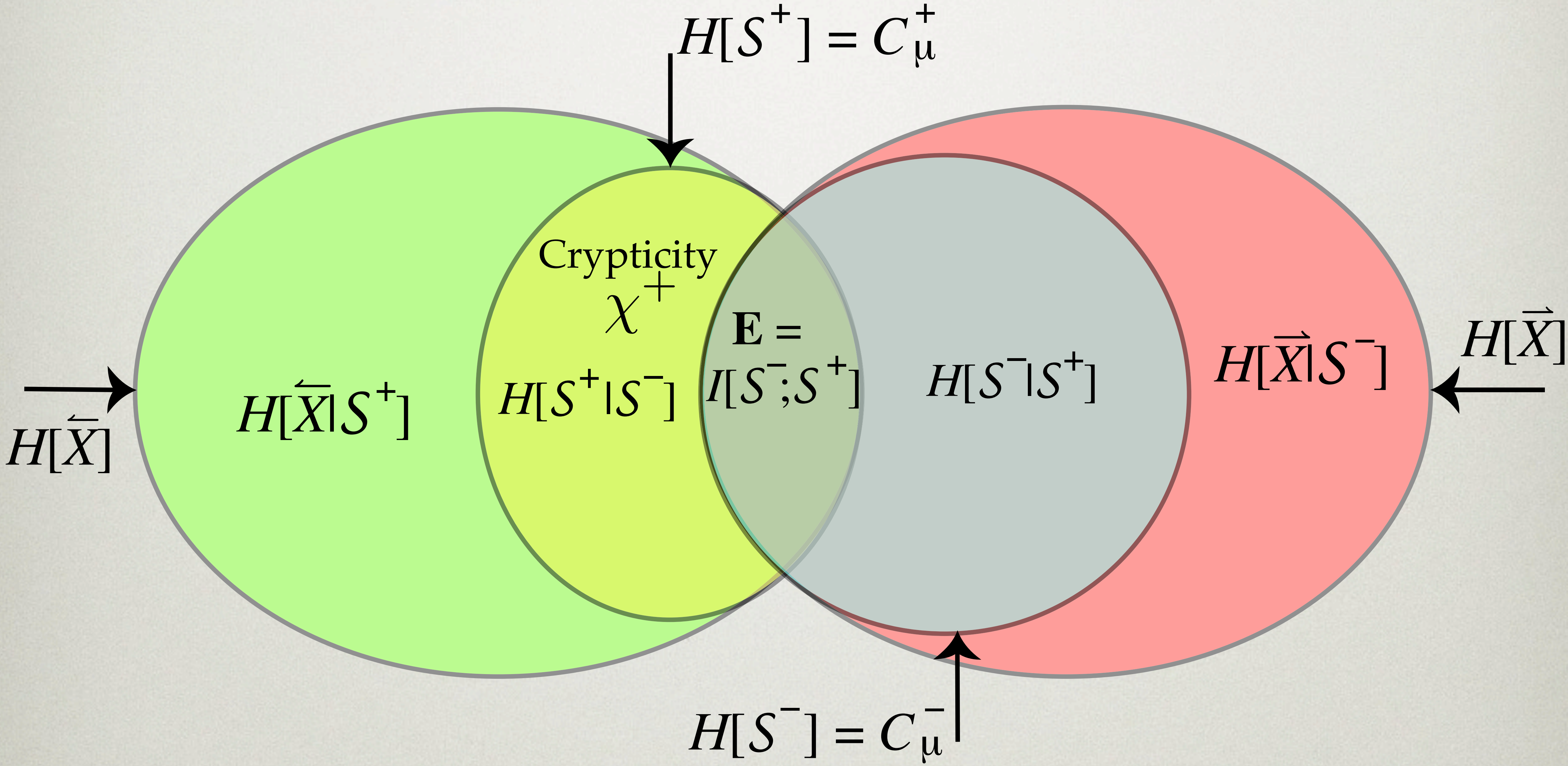
$$H[\mathcal{S}] = C_\mu$$



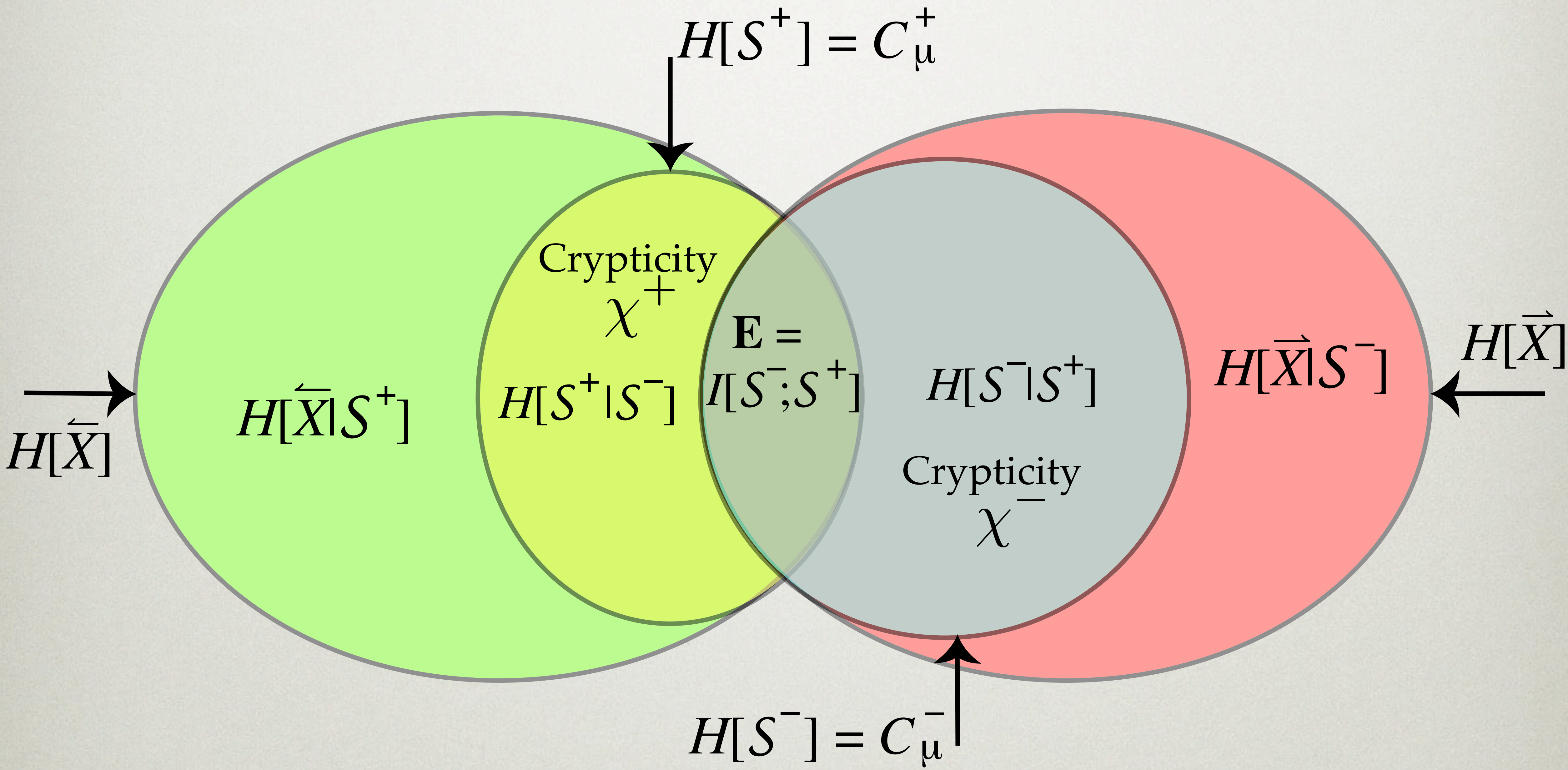
Forward-Reverse ε-machine information Diagram



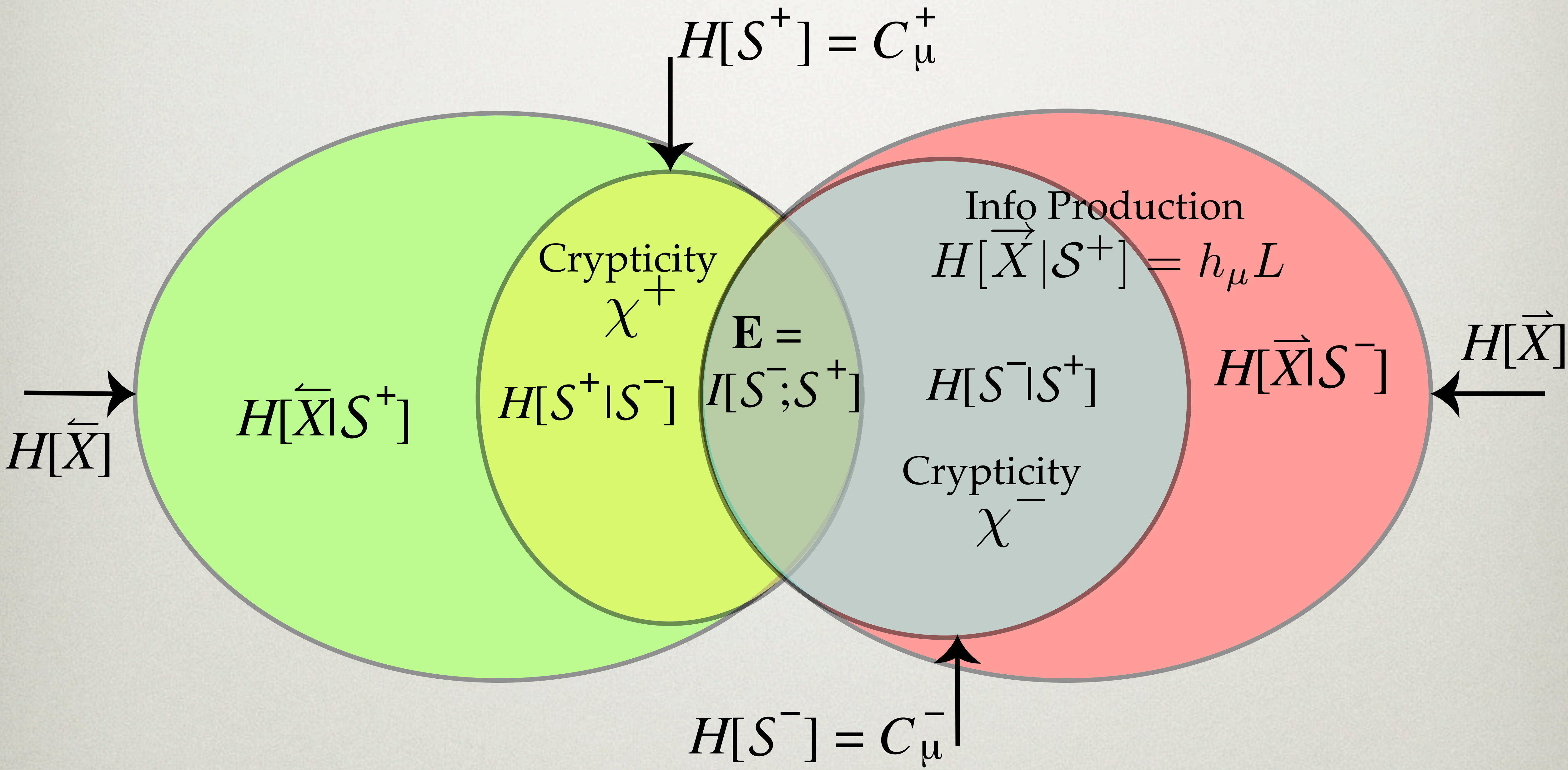
Forward-Reverse ε-machine information Diagram



Forward-Reverse ε-machine information Diagram



Forward-Reverse ε-machine Information Diagram



Forward-Reverse ε-machine information Diagram