Ideal Hypothesis testing and algorithmic information transfer

Bruno Bauwens

Department Systems and Automation
University of Gent

Department Neurology
University Hospital Gent

29/6/2009 Randomness Computability and Logic
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
   - Ratio-test of universal semimeasures
   - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - Total conditional coding theorem
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
Composite Hypothesis tests

Motivation

Procedure

Ratio-test of universal semimeasures

Criterion for universality

Influence tests

Causal semimeasures

Total conditional coding theorem

Incremental coding theorem

Decompositions

Decompositions of algorithmic complexity

Decompositions of mutual information

Randomness tests in hypotheses

Sumtests

Independence tests
Independence and influence in timeseries

Which brain areas communicate?
Improvements

- in algorithms for specific contexts
- in general purpose algorithms

Question: Series of improved general purpose algorithms, what are they converging to? Is there a (non-computable) limit?
Motivation: Algorithms for independence and influence

Improvements

- in algorithms for specific contexts
- in general purpose algorithms

Question: Series of improved general purpose algorithms, what are they converging to? Is there a (non-computable) limit?
Motivation: Algorithms for independence and influence

Improvements

- in algorithms for specific contexts
- in general purpose algorithms

Question: Series of improved general purpose algorithms, what are they converging to? Is there a (non-computable) limit?
### Motivation: Algorithms for independence and influence

**Improvements**
- in algorithms for specific *contexts*
- in general purpose algorithms

**Question:** Series of improved general purpose algorithms, what are they converging to? Is there a (non-computable) limit?
Motivation: Algorithms for independence and influence

Improvements
- in algorithms for specific *contexts*
- in general purpose algorithms

Question: Series of improved general purpose algorithms, what are they converging to? Is there a (non-computable) limit?
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
     - Ratio-test of universal semimeasures
     - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - Total conditional coding theorem
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
Goal of hypothesis tests in science

Discuss inference of probabilistic hypothesis.

- Science = Logic with rules mapped from/to observables → scientists discuss applicability in ...
- Context = Set of observables with some observables fixed
- Hypothesis = set of rules under discussion
- Inference & applicability: experiment -> larger contexts
- Probabilistic rule → observable not consistent in context.
Goal of hypothesis tests in science

Discuss inference of probabilistic hypothesis.

- Science = Logic with rules mapped from/to observables → scientists discuss applicability in ...
- Context = Set of observables with some observables fixed
- Hypothesis = set of rules under discussion
- Inference & applicability: experiment -> larger contexts
- Probabilistic rule → observable not consistent in context.
Goal of hypothesis tests in science

Discuss inference of probabilistic hypothesis.

- Science = Logic with rules mapped from/to observables → scientists discuss applicability in ...
- Context = Set of observables with some observables fixed
- Hypothesis = set of rules under discussion
- Inference & applicability: experiment -> larger contexts
- Probabilistic rule → observable not consistent in context.
Discuss inference of probabilistic hypothesis.

- Science = Logic with *rules* mapped from/to *observables* → scientists *discuss* applicability in ...
- **Context** = Set of *observables* with some *observables* fixed
- **Hypothesis** = set of *rules* under discussion
- Inference & applicability: experiment -> larger contexts
- Probabilistic rule → observable not consistent in context.
Goal of hypothesis tests in science

Discuss inference of probabilistic hypothesis.

- Science = Logic with rules mapped from/to observables → scientists discuss applicability in ...
- Context = Set of observables with some observables fixed
- Hypothesis = set of rules under discussion
- Inference & applicability: experiment -> larger contexts
- Probabilistic rule → observable not consistent in context.
Discuss inference of probabilistic hypothesis.

- Science = Logic with rules mapped from/to observables → scientists discuss applicability in ...
- Context = Set of observables with some observables fixed
- Hypothesis = set of rules under discussion
- Inference & applicability: experiment -> larger contexts
- Probabilistic rule → observable not consistent in context.
Discuss inference of probabilistic hypothesis.

- Science = Logic with rules mapped from/to observables → scientists discuss applicability in ...
- Context = Set of observables with some observables fixed
- Hypothesis = set of rules under discussion
- Inference & applicability: experiment -> larger contexts
- Probabilistic rule → observable not consistent in context.
Procedure

- Hypotheses under discussion: $H^0 \leftrightarrow H^A$
- **Hypothesis test $d$**: function & critical regions
  - If $d(x)$ in critical region: **Reject $H^0$**
    “Either a rare event has occurred or $H^0$ does not describe the data”
  - otherwise: **Fails to reject $H^0$**
- **Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of a some observables in a context
- **Significance $d$ for simple $H^0$** = Probability $H^0$ is **rejected** according to $H^0$
- **Sensitivity $d$ for simple $H^A$** = Probability $H^0$ is **failed to be rejected** according to $H^A$
Procedure

Hypotheses under discussion: $H^0 \leftrightarrow H^A$

- Hypothesis test $d$: function & critical regions
  - If $d(x)$ in critical region: **Reject** $H^0$
    “Either a rare event has occurred or $H^0$ does not describe the data”
  - otherwise: **Fails to reject** $H^0$

- Simple hypothesis = Probabilistic hypothesis fixing probabilities for all observable-values of a some observables in a context

- Significance $d$ for simple $H^0$ = Probability $H^0$ is **rejected** according to $H^0$

- Sensitivity $d$ for simple $H^A$ = Probability $H^0$ is **failed to be rejected** according to $H^A$
Hypotheses under discussion: $H^0 \leftrightarrow H^A$

**Hypothesis test** $d$: function & critical regions
- If $d(x)$ in critical region: **Reject** $H^0$
  “Either a rare event has occurred or $H^0$ does not describe the data”
- Otherwise: **Fails to reject** $H^0$

**Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of a some observables in a context

**Significance $d$ for simple $H^0$** = Probability $H^0$ is rejected according to $H^0$

**Sensitivity $d$ for simple $H^A$** = Probability $H^0$ is failed to be rejected according to $H^A$
Hypotheses under discussion: $H^0 \leftrightarrow H^A$

**Hypothesis test $d$**: function & critical regions
- If $d(x)$ in critical region: **Reject $H^0$**
  - “Either a rare event has occurred or $H^0$ does not describe the data”
- Otherwise: **Fails to reject $H^0$**

**Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of some observables in a context

**Significance $d$ for simple $H^0$** = Probability $H^0$ is **rejected** according to $H^0$

**Sensitivity $d$ for simple $H^A$** = Probability $H^0$ is **failed to be rejected** according to $H^A$
Procedure

- Hypotheses under discussion: $H^0 \leftrightarrow H^A$
- **Hypothesis test $d$**: function & critical regions
  - If $d(x)$ in critical region: **Reject** $H^0$
    “Either a rare event has occurred or $H^0$ does not describe the data”
  - otherwise: **Fails to reject** $H^0$
- **Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of some observables in a context
- **Significance $d$ for simple $H^0$** = Probability $H^0$ is **rejected** according to $H^0$
- **Sensitivity $d$ for simple $H^A$** = Probability $H^0$ is **failed to be rejected** according to $H^A$
Procedure

- Hypotheses under discussion: \( H^0 \leftrightarrow H^A \)
- **Hypothesis test** \( d \): function & critical regions
  - If \( d(x) \) in critical region: **Reject** \( H^0 \)
    - “Either a rare event has occurred or \( H^0 \) does not describe the data”
  - otherwise: **Fails to reject** \( H^0 \)
- **Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of a some observables in a context
- **Significance \( d \) for simple \( H^0 \)** = Probability \( H^0 \) is **rejected** according to \( H^0 \)
- **Sensitivity \( d \) for simple \( H^A \)** = Probability \( H^0 \) is **failed to be rejected** according to \( H^A \)
Hypotheses under discussion: $H^0 \leftrightarrow H^A$

**Hypothesis test $d$**: function & critical regions

- If $d(x)$ in critical region: **Reject $H^0$**
  “Either a rare event has occurred or $H^0$ does not describe the data”
- otherwise: **Fails to reject $H^0$**

**Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of a some observables in a context

- **Significance $d$ for simple $H^0$** = Probability $H^0$ is rejected according to $H^0$
- **Sensitivity $d$ for simple $H^A$** = Probability $H^0$ is failed to be rejected according to $H^A$
Hypotheses under discussion: $H^0 \leftrightarrow H^A$

**Hypothesis test** $d$: function & critical regions
- If $d(x)$ in critical region: **Reject $H^0$**
  “Either a rare event has occurred or $H^0$ does not describe the data”
- otherwise: **Fails to reject** $H^0$

**Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of a some observables in a context

**Significance $d$ for simple $H^0$** = Probability $H^0$ is **rejected** according to $H^0$

**Sensitivity $d$ for simple $H^A$** = Probability $H^0$ is **failed to be rejected** according to $H^A$
**Hypotheses under discussion:** $H^0 \leftrightarrow H^A$

**Hypothesis test $d$:** function & critical regions
- If $d(x)$ in critical region: **Reject $H^0$**
  “Either a rare event has occurred or $H^0$ does not describe the data”
- otherwise: **Fails to reject $H^0$**

**Simple hypothesis** = Probabilistic hypothesis fixing probabilities for all observable-values of a some observables in a context

**Significance $d$ for simple $H^0$** = Probability $H^0$ is **rejected** according to $H^0$

**Sensitivity $d$ for simple $H^A$** = Probability $H^0$ is **failed to be rejected** according to $H^A$
Optimal ratio-test

- One-sided test $d$, critical region $> \lambda$
- Significance and Sensitivity:
  \[ \alpha(\lambda) = \sum \{ P^0(y) : d(y) \geq \lambda \} \]
  \[ \beta(\lambda) = \sum \{ P^A(y) : d(y) \leq \lambda \} \]
- Neyman-Pearson': $\beta \circ \alpha^{-1}$ is uniformly maximal iff $d$ is a.e. monotone function of the ratio-test:
  \[ \frac{P^A(x)}{P^0(x)} \]
- A-priori belief in $H^0, H^A$ is $a^0, a^A$, Bayesian a-posteriori belief:
  \[ \frac{a^0}{a^A} \frac{P^0(x)}{P^A(x)} \]
Optimal ratio-test

- One-sided test $d$, critical region $> \lambda$

Significance en Sensitivity:

$$\alpha(\lambda) = \sum \left\{ P^0(y) : d(y) \geq \lambda \right\}$$

$$\beta(\lambda) = \sum \left\{ P^A(y) : d(y) \leq \lambda \right\}$$

- Neyman-Pearson': $\beta \circ \alpha^{-1}$ is uniformly maximal iff $d$ is a.e. monotone function of the ratio-test:

$$P^A(x)/P^0(x).$$

- A-priori belief in $H^0, H^A$ is $a^0, a^A$, Bayesian a-posteriori belief:

$$a^0 P^0(x)$$

$$\frac{a^A P^A(x)}{a^A P^A(x)}.$$
Optimal ratio-test

- One-sided test $d$, critical region $\geq \lambda$
- Significance en Sensitivity:
  $$\alpha(\lambda) = \sum \{ P^0(y) : d(y) \geq \lambda \}$$
  $$\beta(\lambda) = \sum \{ P^A(y) : d(y) \leq \lambda \}$$
- Neyman-Pearson': $\beta \circ \alpha^{-1}$ is uniformly maximal iff $d$ is a.e. monotone function of the ratio-test:
  $$P^A(x)/P^0(x).$$
- A-priori belief in $H^0, H^A$ is $a^0, a^A$, Bayesian a-posteriori belief:
  $$\frac{a^0 P^0(x)}{a^A P^A(x)}.$$
Optimal ratio-test

- One-sided test $d$, critical region $> \lambda$
- Significance en Sensitivity:
  \[
  \alpha(\lambda) = \sum \{ P^0(y) : d(y) \geq \lambda \}
  \]
  \[
  \beta(\lambda) = \sum \{ P^A(y) : d(y) \leq \lambda \}
  \]
- Neyman-Pearson': $\beta \circ \alpha^{-1}$ is uniformly maximal iff $d$ is a.e. monotone function of the ratio-test:
  \[
  P^A(x)/P^0(x).
  \]
- A-priori belief in $H^0, H^A$ is $a^0, a^A$, Bayesian a-posteriori belief:
  \[
  \frac{a^0}{a^A} \frac{P^0(x)}{P^A(x)}.
  \]
Composite hypothesis tests

- **Composite hypotheses** = set of simple hypotheses
- Different approaches theoretically or empirically optimal in different context for $H^0$, $H^A$ composite:
  1. **Uniformly optimal test**: optimal $\beta \circ \alpha^{-1} \rightarrow$ few cases
  2. **Bayesian approach** $\rightarrow$ subjective
     uniform: Bonferroni-correction
  3. **Generalized maximum likelihood**
     \[
     \frac{\max\{P(x) : P \in H^A\}}{\max\{P(x) : P \in H^0\}}.
     \]
  4. **Information theory**
Composite hypothesis tests

- **Composite hypotheses** = set of simple hypotheses
- Different approaches theoretically or empirically optimal in different context for $H^0, H^A$ composite:
  1. **Uniformly optimal test**: optimal $\beta \circ \alpha^{-1} \rightarrow$ few cases
  2. **Bayesian approach** $\rightarrow$ subjective
     - uniform: Bonferroni-correction
  3. **Generalized maximum likelihood**
     \[
     \frac{\max\{P(x) : P \in H^A\}}{\max\{P(x) : P \in H^0\}}.
     \]
  4. **Information theory**

Bruno Bauwens  Hypotheses and AIT
Composite hypothesis tests

- **Composite hypotheses** = set of simple hypotheses
- Different approaches theoretically or empirically optimal in different context for $H^0$, $H^A$ composite:
  1. *Uniformly optimal test*: optimal $\beta \circ \alpha^{-1} \rightarrow$ few cases
  2. *Bayesian approach* $\rightarrow$ subjective
     - uniform: Bonferroni-correction
  3. *Generalized maximum likelihood*
     \[
     \max \{ P(x) : P \in H^A \} \\
     \max \{ P(x) : P \in H^0 \}.
     \]
  4. *Information theory*

Hypotheses and AIT
**Composite hypothesis tests**

**Composite hypotheses** = set of simple hypotheses

Different approaches theoretically or empirically optimal in different context for $H^0$, $H^A$ composite:

1. **Uniformly optimal test**: optimal $\beta \circ \alpha^{-1} \rightarrow$ few cases
2. **Bayesian approach** $\rightarrow$ subjective
   - uniform: Bonferroni-correction
3. **Generalized maximum likelihood**
   $$\frac{\max\{P(x) : P \in H^A\}}{\max\{P(x) : P \in H^0\}}.$$  
4. **Information theory**

**Motivation**

**Procedure**

**Ratio-test of universal semimeasures**

**Criterion for universality**
**Composite hypothesis tests**

- **Composite hypotheses** = set of simple hypotheses
- Different approaches theoretically or empirically optimal in different context for $H^0$, $H^A$ composite:
  1. *Uniformly optimal test*: optimal $\beta \circ \alpha^{-1} \to$ few cases
  2. *Bayesian approach* $\to$ subjective
     - uniform: Bonferroni-correction
  3. *Generalized maximum likelihood*
     \[
     \frac{\max\{P(x) : P \in H^A\}}{\max\{P(x) : P \in H^0\}}.
     \]
  4. *Information theory*
**Composite hypothesis tests**

- **Composite hypotheses** = set of simple hypotheses
- Different approaches theoretically or empirically optimal in different context for $H^0, H^A$ composite:
  1. **Uniformly optimal test**: optimal $\beta \circ \alpha^{-1} \rightarrow$ few cases
  2. **Bayesian approach** $\rightarrow$ subjective
     - uniform: Bonferroni-correction
  3. **Generalized maximum likelihood**
     \[
     \frac{\max\{P(x) : P \in H^A\}}{\max\{P(x) : P \in H^0\}}.
     \]

- **Information theory**

---

**Bruno Bauwens**

Hypotheses and AIT
Composite hypothesis tests

- Composite hypotheses = set of simple hypotheses
- Different approaches theoretically or empirically optimal in different context for $H^0$, $H^A$ composite:
  1. **Uniformly optimal test**: optimal $\beta \circ \alpha^{-1} \rightarrow \text{few cases}$
  2. **Bayesian approach** → subjective
     uniform: Bonferroni-correction
  3. **Generalized maximum likelihood**
     \[
     \frac{\max\{P(x) : P \in H^A\}}{\max\{P(x) : P \in H^0\}}.
     \]
  4. **Information theory**
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
   - Ratio-test of universal semimeasures
   - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - Total conditional coding theorem
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
Length conditional semimeasures

\( P \) is length conditional iff

\[
\forall n : \sum \{ P(x) : x \in 2^n \} \leq 1
\]

- No nice coding result for monotones semimeasures (Gacs, previous talk)
- For most experiments length of the measurements contains very little information

“length conditional semimeasure” abbr. to “semimeasure”
Length conditional semimeasures

\( P \) is length conditional iff

\[
\forall n : \sum \{ P(x) : x \in 2^n \} \leq 1
\]

- No nice coding result for monotones semimeasures (Gacs, previous talk)
  - For most experiments length of the measurements contains very little information
  - “length conditional semimeasure” abbr. to “semimeasure”
Length conditional semimeasures

$P$ is length conditional iff

$$\forall n : \sum \{ P(x) : x \in 2^n \} \leq 1$$

- No nice coding result for monotones semimeasures (Gacs, previous talk)
- For most experiments length of the measurements contains very little information

“length conditional semimeasure” abbr. to “semimeasure”
Length conditional semimeasures

\[ P \text{ is length conditional iff } \]

\[ \forall n : \sum \{ P(x) : x \in 2^n \} \leq 1 \]

- No nice coding result for monotones semimeasures (Gacs, previous talk)
- For most experiments length of the measurements contains very little information
  “length conditional semimeasure” abbr. to “semimeasure”
### Universal semimeasures

#### Definition

Semimeasures $P, Q,$

- $P$ dominates $Q$ ($P \geq^* Q$) iff $\exists c \forall x : P(x) \geq cQ(x)$
- $P =^* Q$ iff $P \leq^* Q$ and $P \geq^* Q$
- $m$ is universal in $H$ iff $\forall P \in H : P \leq^* m$
- Ratio test of $H^0, H^A$:

$$d(x) = m^A(x)/m^0(x),$$

- Generalized maximum likelihood testing (up to *-constant)
- Enumerable cases: Bayesian approach $m^A =^* \sum a_i P_i$
Universal semimeasures

Definition

Semimeasures $P, Q,$
- $P$ dominates $Q$ ($P \succeq^* Q$) iff $\exists c \forall x : P(x) \geq cQ(x)$
- $P =^* Q$ iff $P \preceq^* Q$ and $P \succeq^* Q$
- $m$ is universal in $H$ iff $\forall P \in H: P \preceq^* m$
- Ratio test of $H^0, H^A$:
  $$d(x) = m^A(x)/m^0(x),$$

- Generalized maximum likelihood testing (up to *-constant)
- Enumerable cases: Bayesian approach $m^A =^* \sum a_i P_i$
Universal semimeasures

**Definition**

Semimeasures $P, Q$,

- $P$ dominates $Q$ ($P \geq^* Q$) iff $\exists c \forall x : P(x) \geq cQ(x)$
- $P =^* Q$ iff $P \leq^* Q$ and $P \geq^* Q$
- $m$ is universal in $H$ iff $\forall P \in H: P \leq^* m$

Ratio test of $H^0, H^A$:

$$d(x) = m^A(x)/m^0(x),$$

- Generalized maximum likelihood testing (up to $^*$-constant)
- Enumerative cases: Bayesian approach $m^A =^* \sum a_i P_i$
Universal semimeasures

Definition

Semimeasures \( P, Q, \)
- \( P \text{ dominates } Q \) (\( P \geq^* Q \)) iff \( \exists c \forall x : P(x) \geq cQ(x) \)
- \( P =^* Q \) iff \( P \leq^* Q \) and \( P \geq^* Q \)
- \( m \) is universal in \( H \) iff \( \forall P \in H : P \leq^* m \)
- Ratio test of \( H^0, H^A \):
  \[
  d(x) = m^A(x)/m^0(x),
  \]

- Generalized maximum likelihood testing (up to \(^*-\)constant)
- Enumerable cases: Bayesian approach \( m^A =^* \sum a_i P_i \)
Universal semimeasures

Definition

Semimeasures $P$, $Q$,

- *$P$ dominates $Q$ ($P \succeq^* Q$) iff $\exists c \forall x : P(x) \geq cQ(x)$*
- *$P =^* Q$ iff $P \preceq^* Q$ and $P \succeq^* Q$*
- *$m$ is universal in $H$ iff $\forall P \in H : P \preceq^* m$*
- *Ratio test of $H^0$, $H^A$:

$$d(x) = \frac{m^A(x)}{m^0(x)},$$

- Generalized maximum likelihood testing (up to *-constant)
- Enumerable cases: Bayesian approach $m^A =^* \sum a_i P_i$
Universal semimeasures

**Definition**

Semimeasures $P, Q,$

- $P$ dominates $Q$ ($P \succeq^* Q$) iff $\exists c \forall x : P(x) \geq cQ(x)$
- $P =^* Q$ iff $P \preceq^* Q$ and $P \succeq^* Q$
- $m$ is universal in $H$ iff $\forall P \in H : P \preceq^* m$
- Ratio test of $H^0, H^A$:

\[ d(x) = \frac{m^A(x)}{m^0(x)}, \]

- Generalized maximum likelihood testing (up to $^*$-constant)
-Enumerable cases: Bayesian approach $m^A =^* \sum a_i P_i$
Interpretation

For $H^0$ enumerable, for $i < k/(\log k)^2$:

$$\frac{P_i^0(x)}{m^A(x)} \leq^* \frac{km^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:
- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical 'significance' is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore: significance must be defined empirically in context.
For $H^0$ enumerable, for $i < k/(\log k)^2$:

$$\frac{P_i^0(x)}{m^A(x)} \leq^* \frac{km^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:

- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical 'significance' is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore: significance must be defined empirically in context.
Interpretation

For $H^0$ enumerable, for $i < k/(\log k)^2$:

$$\frac{P_i^0(x)}{m^A(x)} \leq * \frac{km_i^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:

- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical 'significance' is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore: significance must be defined empirically in context.
Interpretation

For $H^0$ enumerable, for $i < k/(\log k)^2$:

$$\frac{P_i^0(x)}{m^A(x)} \leq^* \frac{km^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:

- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical 'significance' is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore: significance must be defined empirically in context.
Interpretation

For $H^0$ enumerable, for $i < k/\log(k)^2$:

$$\frac{P_i^0(x)}{m^A(x)} \leq^* \frac{km^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:

- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical ‘significance’ is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore: significance must be defined empirically in context.
For $H^0$ enumerable, for $i < k/(\log k)^2$:

$$\frac{P^0_i(x)}{m^A(x)} \leq * \frac{km^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:
- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical 'significance' is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore:
  significance must be defined empirically in context.
For $H^0$ enumerable, for $i < k/(\log k)^2$:

$$\frac{P^0_i(x)}{m^A(x)} \leq^* \frac{km^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:

- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical 'significance' is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore: significance must be defined empirically in context.
Interpretation

For $H^0$ enumerable, for $i < k/(\log k)^2$:

$$\frac{P_i^0(x)}{m^A(x)} \leq * \frac{km^0(x)}{m^A(x)}$$

A high $d(x)$ means that either:

- A complex model from the zero hypothesis describes data $x$
- The alternate hypothesis $m^A$ better describes data $x$
- A rare event has occurred

→ Theoretical 'significance' is lower than $1/d(x)$
→ In practice, (ex. causality, independence), the first interpretation cannot be completely eliminated, therefore: significance must be defined empirically in context.
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
   - Ratio-test of universal semimeasures
   - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - Total conditional coding theorem
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
Let $S$ be a set of enumerable semimeasures

- $S$ is *testable* iff for some computable logic expression $L$:
  \[ P \in S \iff \forall t, n \leq t : L(P^n_t). \]
  with $P^n_t$ = finite restriction of $P_t$ on $2^n$

- $S$ is *convex* iff $\forall P, Q \in S, a \in [0, 1] : aP + (1 - a)Q \in R$

- $S$ has a *computable monotone convex upper-bounded* iff a computable $R$ exists, such that for all $a, P, Q, P', Q'$
  \[ aP + (1 - a)Q \preceq R(P, Q, a) \]

- $P' \succeq P; Q' \succeq Q \Rightarrow R(P', Q', a) \succeq R(P, Q, a)$

Remark: $S$ is convex, implies $S$ is *computable monotone convex upper-bounded.*
Let $S$ be a set of enumerable semimeasures

- $S$ is *testable* iff for some computable logic expression $L$:

$$P \in S \iff \forall t, n \leq t : L(P^n_t).$$

with $P^n_t =$ finite restriction of $P_t$ on $2^n$

- $S$ is *convex* iff $\forall P, Q \in S, a \in [0, 1] : aP + (1 - a)Q \in R$

- $S$ has a *computable monotone convex upper-bounded* iff a computable $R$ exists, such that for all $a, P, Q, P', Q'$

$$aP + (1 - a)Q \leq R(P, Q, a)$$

$$P' \succeq P; Q' \succeq Q \Rightarrow R(P', Q', a) \succeq R(P, Q, a)$$

Remark: $S$ is convex, implies $S$ is *computable monotone convex upper-bounded.*
Let $S$ be a set of enumerable semimeasures

- $S$ is *testable* iff for some computable logic expression $L$:

$$P \in S \iff \forall t, n \leq t : L(P^n_t).$$

with $P^n_t$ = finite restriction of $P_t$ on $2^n$

- $S$ is *convex* iff $\forall P, Q \in S, a \in [0, 1] : aP + (1 - a)Q \in R$

- $S$ has a *computable monotone convex upper-bounded* iff a computable $R$ exists, such that for all $a, P, Q, P', Q'$

$$aP + (1 - a)Q \leq R(P, Q, a)$$

$$P' \geq P; Q' \geq Q \Rightarrow R(P', Q', a) \geq R(P, Q, a)$$

Remark: $S$ is *convex*, implies $S$ is *computable monotone convex upper-bounded.*
Let $S$ be a set of enumerable semimeasures

- $S$ is *testable* iff for some computable logic expression $L$:
  \[ P \in S \iff \forall t, n \leq t : L(P_t^n). \]
  with $P_t^n$ = finite restriction of $P_t$ on $2^n$

- $S$ is *convex* iff $\forall P, Q \in S, a \in [0, 1] : aP + (1 - a)Q \in R$

- $S$ has a *computable monotone convex upper-bounded* iff a computable $R$ exists, such that for all $a, P, Q, P', Q'$
  \[ aP + (1 - a)Q \leq R(P, Q, a) \]
  \[ P' \geq P; Q' \geq Q \Rightarrow R(P', Q', a) \geq R(P, Q, a) \]

Remark: $S$ is convex, implies $S$ is computable monotone convex upper-bounded.
Let $S$ be a set of enumerable semimeasures
- $S$ is *testable* iff for some computable logic expression $L$:
  $$P \in S \iff \forall t, n \leq t : L(P^n_t).$$
  with $P^n_t$ = finite restriction of $P_t$ on $2^n$
- $S$ is *convex* iff $\forall P, Q \in S, a \in [0, 1] : aP + (1 - a)Q \in R$
- $S$ has a *computable monotone convex upper-bounded* iff a computable $R$ exists, such that for all $a, P, Q, P', Q'$
  $$aP + (1 - a)Q \leq R(P, Q, a)$$

$$P' \geq P; Q' \geq Q \Rightarrow R(P', Q', a) \geq R(P, Q, a)$$

Remark: $S$ is *convex*, implies $S$ is *computable monotone convex upper-bounded*. 

Bruno Bauwens  
Hypotheses and AIT
Proposition

\[ \Sigma = \text{set of enumerable semimeasures} \]

(i) If \( S \) is testable, \( P^{(0)} = 0 \in S \), than \( S \cap \Sigma \) is enumerable.

(ii) If \( S \) has a computable monotone convex upper-bounded, and \( S \cap \Sigma \) is enumerable, than it has a universal element.
Proposition

\[ \Sigma = \text{set of enumerable semimeasures} \]

(i) If \( S \) is testable, \( P^{(0)} = 0 \in S \), than \( S \cap \Sigma \) is enumerable.

(ii) If \( S \) has a computable monotone convex upper-bounded, and \( S \cap \Sigma \) is enumerable, than it has a universal element.
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, ...$
- Mixture

$$m^S(x) = \sum a_i P_i$$

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:
- Semimeasures
- Conditional semimeasures
- Uniform conditional semimeasures (further)
- Independent semimeasures: $P(x, y) = Q(x)R(y)$
- Conditional causal semimeasures (further)
- Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- **Enumeration**: $S = P_0, P_1, \ldots$
  - **Mixture**
    \[ m^S(x) = \sum a_i P_i \]
  
  → Very difficult to approximate

Hypotheses with universal enumerable semimeasure:
- **Semimeasures**
- **Conditional semimeasures**
- **Uniform conditional semimeasures (further)**
- **Independent semimeasures**: $P(x, y) = Q(x)R(y)$
- **Conditional causal semimeasures (further)**
- **Total conditional causal semimeasures (further)**
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, ...$
- Mixture

$$m^S(x) = \sum a_i P_i$$

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:
- Semimeasures
- Conditional semimeasures
- Uniform conditional semimeasures (further)
- Independent semimeasures: $P(x, y) = Q(x)R(y)$
- Conditional causal semimeasures (further)
- Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, \ldots$
- Mixture
  \[ m^S(x) = \sum a_i P_i \]

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:
- Semimeasures
- Conditional semimeasures
- Uniform conditional semimeasures (further)
- Independent semimeasures: $P(x, y) = Q(x)R(y)$
- Conditional causal semimeasures (further)
- Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, ...$
- Mixture

$$m^S(x) = \sum a_i P_i$$

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:

- Semimeasures
- Conditional semimeasures
- Uniform conditional semimeasures (further)
- Independent semimeasures: $P(x, y) = Q(x)R(y)$
- Conditional causal semimeasures (further)
- Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, \ldots$
- Mixture

$$m^S(x) = \sum a_i P_i$$

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:

- Semimeasures
  - Conditional semimeasures
  - Uniform conditional semimeasures (further)
  - Independent semimeasures: $P(x, y) = Q(x)R(y)$
  - Conditional causal semimeasures (further)
  - Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, ...$
- Mixture

$$m^S(x) = \sum a_i P_i$$

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:
- Semimeasures
- Conditional semimeasures
  - Uniform conditional semimeasures (further)
  - Independent semimeasures: $P(x, y) = Q(x)R(y)$
  - Conditional causal semimeasures (further)
  - Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, \ldots$
- Mixture
  \[ m^S(x) = \sum a_i P_i \]

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:
  - Semimeasures
  - Conditional semimeasures
  - Uniform conditional semimeasures (further)
    - Independent semimeasures: $P(x, y) = Q(x)R(y)$
    - Conditional causal semimeasures (further)
    - Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, \ldots$
- Mixture
  $$m^S(x) = \sum a_i P_i$$

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:
  - Semimeasures
  - Conditional semimeasures
  - Uniform conditional semimeasures (further)
  - Independent semimeasures: $P(x, y) = Q(x)R(y)$
  - Conditional causal semimeasures (further)
  - Total conditional causal semimeasures (further)
Explicit construction of universal element of $S$

- Enumeration: $S = P_0, P_1, ...$
- Mixture
  
  $$m^S(x) = \sum a_i P_i$$

→ Very difficult to approximate

Hypotheses with universal enumerable semimeasure:

- Semimeasures
- Conditional semimeasures
- Uniform conditional semimeasures (further)
- Independent semimeasures: $P(x, y) = Q(x)R(y)$
- Conditional causal semimeasures (further)
- Total conditional causal semimeasures (further)
Outline

1 Composite Hypothesis tests
   • Motivation
   • Procedure
   • Ratio-test of universal semimeasures
   • Criterion for universality

2 Influence tests
   • Causal semimeasures
   • Total conditional coding theorem
   • Incremental coding theorem

3 Decompositions
   • Decompositions of algorithmic complexity
   • Decompositions of mutual information

4 Randomness tests in hypotheses
   • Sumtests
   • Independence tests
From now on: \( x, y, z \in 2^n \)

\[
x^i = x_1 x_2 \ldots x_i
\]

If \( v \in 2^i \) then

\[
P(v | y) = \sum \{ P(vw) : w \in 2^{n-i} \}
\]

\[
P(x^i, y^j) = \sum \{ P(x^i v, y^j w) : v \in 2^{n-i}, w \in 2^{n-j} \}
\]

A conditional semimeasure \( P \) is **uniform conditional** \( (P(x | y)) \) iff

\[
\forall y, z : P(\epsilon | y) = P(\epsilon | z)
\]
Notation and Definitions (1)

From now on: \( x, y, z \in 2^n \)
\( x^i = x_1 x_2 \ldots x_i \)
If \( v \in 2^i \) then \( P(v|y) = \sum \{ P(vw) : w \in 2^{n-i} \} \)
\( P(x^i, y^j) = \sum \{ P(x^i v, y^j w) : v \in 2^{n-i}, w \in 2^{n-j} \} \)

A conditional semimeasure \( P \) is **uniform conditional** \((P(x|y))\) iff
\[ \forall y, z : P(\epsilon|y) = P(\epsilon|z) \]
From now on: $x, y, z \in 2^n$

$x^i = x_1 x_2 \ldots x_i$

If $v \in 2^i$ then $P(v|y) = \sum \{P(vw) : w \in 2^{n-i}\}$

$P(x^i, y^j) = \sum \{P(x^i v, y^j w) : v \in 2^{n-i}, w \in 2^{n-j}\}$

- A conditional semimeasure $P$ is \textit{uniform conditional} $(P(x|y))$ iff

\[ \forall y, z : P(\epsilon|y) = P(\epsilon|z) \]
From now on: \( x, y, z \in 2^n \)

\( x^i = x_1 x_2 \ldots x_i \)

If \( v \in 2^i \) then 
\[
P(v|y) = \sum \{ P(vw) : w \in 2^{n-i} \}
\]

\[
P(x^i, y^j) = \sum \{ P(x^i v, y^j w) : v \in 2^{n-i}, w \in 2^{n-j} \}
\]

- A conditional semimeasure \( P \) is **uniform conditional** \((P(x|y))\) iff
\[
\forall y, z : P(\epsilon|y) = P(\epsilon|z)
\]
A conditional semimeasure \( P \) is *conditional causal* \( (P(x|y↑)) \) iff

\[
\forall i, x, y, z : y^i = z^i \Rightarrow P(x^{i+1}|y) = P(x^{i+1}|z)
\]

The *associated causal semimeasure* \( P(x|y↑) \) associated with a semimeasure \( P(x,y) \) is:

\[
P(x|y↑) = \frac{P(x^1, y^0) P(x^2, y^1) \cdots P(x^n, y^{n-1})}{P(x^0, y^0) P(x^1, y^1) \cdots P(x^{n-1}, y^{n-1})}
\]

A conditional semimeasure \( P(x|y) \) is *associated causal* \( (P(x|y↑)) \) if it is associated with an enumerable semimeasure \( P(x,y) \).
A conditional semimeasure $P$ is \textit{conditional causal} $(P(x|y \uparrow))$ iff
\[ \forall i, x, y, z : y^i = z^i \Rightarrow P(x^{i+1}|y) = P(x^{i+1}|z) \]

The \textit{associated causal semimeasure} $P(x|y \uparrow)$ \textit{associated with a semimeasure} $P(x, y)$ is:
\[ P(x|y \uparrow) = \frac{P(x^1, y^0) P(x^2, y^1) \cdots P(x^n, y^{n-1})}{P(x^0, y^0) P(x^1, y^1) \cdots P(x^{n-1}, y^{n-1})} \]

A conditional semimeasure $P(x|y)$ is \textit{associated causal} $(P(x|y \uparrow))$ if it is associated with an enumerable semimeasure $P(x, y)$. 
A conditional semimeasure $P$ is \textit{conditional causal} \((P(x \mid y ↑))\) \textit{iff}

\[
\forall i, x, y, z : y^i = z^i \Rightarrow P(x^{i+1} \mid y) = P(x^{i+1} \mid z)
\]

The \textit{associated causal semimeasure} $P(x \mid y ↑)$ \textit{associated with a semimeasure} $P(x, y)$ is:

\[
P(x \mid y ↑) = \frac{P(x^1, y^0)}{P(x^0, y^0)} \cdot \frac{P(x^2, y^1)}{P(x^1, y^1)} \cdot \ldots \cdot \frac{P(x^n, y^{n-1})}{P(x^{n-1}, y^{n-1})}
\]

A conditional semimeasure $P(x \mid y)$ is \textit{associated causal} \((P(x \mid y ↑))\) \textit{if it is associated with an enumerable semimeasure} $P(x, y)$. 
Relations between conditional and causal semimeasures

- Conditional: $m(x|y)$
- Uniform conditional: $m(x \mid y)$
- Associated causal
- Associated with an universal $m(x,y)$
- Conditional causal: $m(x \mid y \uparrow)$
Unstable inference of influence

**Proposition**

For enumerable $P$, with $P(x, y) \geq 2^{-2n}$, there are enumerable $Q, R$ with $P =^* Q =^* R$ and

$$\log \frac{Q(x|y \uparrow)}{P(x|y \uparrow)} > o(n) \quad \text{and} \quad \log \frac{P(x|y \uparrow)}{R(x|y \uparrow)} > o(n).$$

**Corollary**

The set of associated causal semimeasures associated with $P(x, y) \geq 2^{-2n}$, has no universal element.
Proposition

For enumerable $P$, with $P(x, y) \geq 2^{-2n}$, there are enumerable $Q, R$ with $P = * Q = * R$ and

$$\log \frac{Q(x | y \uparrow)}{P(x | y \uparrow)} > o(n) \quad \log \frac{P(x | y \uparrow)}{R(x | y \uparrow)} > o(n).$$

Corollary

The set of associated causal semimeasures associated with $P(x, y) \geq 2^{-2n}$, has no universal element.
The ratio-test:

\[
\frac{m(x|y)}{m(x|y \uparrow)}
\]

depends on the choice of \( m \).

- Connection to observed instability of influence measure from Huffman trees?
- Inference of influence with ideal compression using \( K \) (coding theorem) either:
  - needs further restrictions on \( K \)
  - is unstable
The ratio-test:

\[ \frac{m(x|y)}{m(x|y \uparrow)} \]

depends on the choice of \( m \).

Connection to observed instability of influence measure from Huffman trees?

Inference of influence with ideal compression using \( K \) (coding theorem) either:
- needs further restrictions on \( K \)
- is unstable
The ratio-test:

\[
\frac{m(x|y)}{m(x|y \uparrow)}
\]

depends on the choice of \( m \).

Connection to observed instability of influence measure from Huffman trees?

Inference of influence with ideal compression using \( K \) (coding theorem) either:
- needs further restrictions on \( K \)
- is unstable
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
   - Ratio-test of universal semimeasures
   - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - **Total conditional coding theorem**
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
Theorem (Coding)

\[ K(x) =^{+} \log m(x) \]
\[ K(x|y) =^{+} \log m(x|y) \]

Apply search-heuristics for data compression to estimate \( m(x), m(x|y) \)
Composite Hypothesis tests
Influence tests
Decompositions
Randomness tests in hypotheses

(Length conditional prefix-free) Kolmogorov complexity

\( \langle ., . \rangle \rightarrow \) computable bijective pairing function.

\[
K_t(x) = \min \{ l(p) : \Phi_t(p, n) \downarrow = x \} \\
K(x) = \lim_{t \to \infty} K_t(x) \\
K(x, y) = K(\langle x, y \rangle) \\
K(x|y) = \min \{ l(p) : \Phi(p, y, n) \downarrow = x \}
\]

**Definition**

\[
K(x|y) = \min \{ l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \}
\]
(Length conditional prefix-free) Kolmogorov complexity

\langle ., . \rangle \rightarrow \text{computable bijective pairing function.}

\[
K_t(x) = \min \{ l(p) : \Phi_t(p, n) \downarrow = x \}
\]

\[
K(x) = \lim_{t \to \infty} K_t(x)
\]

\[
K(x, y) = K(\langle x, y \rangle)
\]

\[
K(x|y) = \min \{ l(p) : \Phi(p, y, n) \downarrow = x \}
\]

**Definition**

\[
K(x|y) = \min \{ l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \}
\]
Composite Hypothesis tests
Influence tests
Decompositions
Randomness tests in hypotheses

Causal semimeasures
Total conditional coding theorem
Incremental coding theorem

(Length conditional prefix-free) Kolmogorov complexity

\langle . , . \rangle \rightarrow \text{computable bijective pairing function.}

\begin{align*}
K_t(x) &= \min \{ l(p) : \Phi_t(p, n) \downarrow = x \} \\
K(x) &= \lim_{t \to \infty} K_t(x) \\
K(x, y) &= K(\langle x, y \rangle) \\
K(x|y) &= \min \{ l(p) : \Phi(p, y, n) \downarrow = x \}
\end{align*}

Definition

\[ K(x|y) = \min \{ l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \} \]
\langle ., . \rangle \rightarrow \text{computable bijective pairing function.}

\begin{align*}
K_t(x) & = \min \{l(p) : \Phi_t(p, n) \downarrow = x\} \\
K(x) & = \lim_{t \rightarrow \infty} K_t(x) \\
K(x, y) & = K(\langle x, y \rangle) \\
K(x|y) & = \min \{l(p) : \Phi(p, y, n) \downarrow = x\}
\end{align*}

\textbf{Definition}

\[ K(x|y) = \min \{l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \} \]
Difference total conditional and conditional complexity

Proposition

For all $n$ there are $x, y \in 2^n$ with:

$$K(x \| y) - K(x | y) \geq^+ n$$

Difference due to Halting information

Proposition

$$K(x \| y) - K(x | y) \leq^+ K(y) - K'(y) + O(\log k_{xy})$$
Proposition

For all \( n \) there are \( x, y \in 2^n \) with:

\[
K(x|y) - K(x|y) \geq^+ n
\]

Difference due to Halting information

Proposition

\[
K(x|y) - K(x|y) \leq^+ K(y) - K'(y) + O(\log k_{xy})
\]
Proposition

For all $n$ there are $x, y \in 2^n$ with:

$$K(x|y) - K(x|y) \geq n$$

Difference due to Halting information

Proposition

$$K(x|y) - K(x|y) \leq K(y) - K'(y) + O(\log k_{xy})$$
Proposition

For all \( n \) there are \( x, y \in 2^n \) with:

\[
K(x \mid y) - K(x \mid y) \geq^+ n
\]

Difference due to Halting information

Proposition

\[
K(x \mid y) - K(x \mid y) \leq^+ K(y) - K'(y) + O(\log k_{xy})
\]
Definition

\[ t_k = \min \{ t : m(\epsilon) - m_t(\epsilon) \leq 2^{-k} \} \]

\[ k_{xy} = \min \{ k : K_{t_k}(x, y|n^*) =^+ K(x, y|n^*) \} \]

Lemma

Let \( m^i, i = 0, 1 \) be two universal semimeasures
Let \( k^i_{xy} \) be the corresponding \( m^i \)-depths, then:

\[ k^0_{xy} = k^1_{xy} \pm O(\log k^1_{xy}) \]
**Definition**

\[ t_k = \min \{ t : m(\epsilon) - m_t(\epsilon) \leq 2^{-k} \} \]

\[ k_{xy} = \min \{ k : K_{t_k}(x, y|n^*) =^+ K(x, y|n^*) \} \]

**Lemma**

Let \( m^i, i = 0, 1 \) be two universal semimeasures
Let \( k^i_{xy} \) be the corresponding \( m^i \)-depths, then:

\[ k^0_{xy} = k^1_{xy} \pm O(\log k^1_{xy}) \]
**m-depth**

**Definition**

\[ t_k = \min\{ t : m(\epsilon) - m_t(\epsilon) \leq 2^{-k} \} \]

\[ k_{xy} = \min\{ k : K_{t_k}(x, y|n^*) =^+ K(x, y|n^*) \} \]

**Lemma**

Let \( m_i, i = 0, 1 \) be two universal semimeasures. Let \( k_{xy}^i \) be the corresponding \( m^i \)-depths, then:

\[ k_{xy}^0 = k_{xy}^1 \pm O(\log k_{xy}^1) \]
**Definition**

\[ t_k = \min \{ t : m(\epsilon) - m_t(\epsilon) \leq 2^{-k} \} \]

\[ k_{xy} = \min \{ k : K_{t_k}(x, y|n^*) =^+ K(x, y|n^*) \} \]

**Lemma**

Let \( m^i, i = 0, 1 \) be two universal semimeasures

Let \( k_{xy}^i \) be the corresponding \( m^i \)-depths, then:

\[ k_{xy}^0 = k_{xy}^1 \pm O(\log k_{xy}^1) \]
**Definition**

\[ t_k = \min \{ t : m(\epsilon) - m_t(\epsilon) \leq 2^{-k} \} \]

\[ k_{xy} = \min \{ k : K_{t_k}(x, y|n^*) =^+ K(x, y|n^*) \} \]

**Lemma**

*Let \( m^i, i = 0, 1 \) be two universal semimeasures. Let \( k^i_{xy} \) be the corresponding \( m^i \)-depths, then:*

\[ k^0_{xy} = k^1_{xy} \pm O(\log k^1_{xy}) \]
Proposition (Total coding theorem)

\[ K(x \mid y) = \log m(x \mid y) \pm O(\log k_{xy}) \]
Outline

1 Composite Hypothesis tests
   • Motivation
   • Procedure
   • Ratio-test of universal semimeasures
   • Criterion for universality

2 Influence tests
   • Causal semimeasures
   • Total conditional coding theorem
   • Incremental coding theorem

3 Decompositions
   • Decompositions of algorithmic complexity
   • Decompositions of mutual information

4 Randomness tests in hypotheses
   • Sumtests
   • Independence tests
Incremental computation

\[ x^i = x_1 \ldots x_i \]
\[ \Phi(p, x \uparrow) \downarrow = y \text{ iff } \forall i < n : \Phi(p, x^i, n) = \downarrow y_{i+1} \]
Incremental computation

\[ x^i = x_1 \ldots x_i \]
\[ \Phi(p, x \uparrow) \downarrow = y \text{ iff } \forall i < n : \Phi(p, x^i, n) = \downarrow y_{i+1} \]

\begin{align*}
  x & \quad y \\
  = & \quad = \\
  x_4 & \quad y_4 \\
  x_3 & \quad y_3 \\
  x_2 & \quad y_2 \\
  x_1 & \quad y_1
\end{align*}
Incremental computation

\[ x^i = x_1 \ldots x_i \]
\[ \Phi(p, x \uparrow) \downarrow = y \text{ iff } \forall i < n : \Phi(p, x^i, n) = \downarrow y_{i+1} \]

\[
\begin{array}{cc}
x & y \\
= & = \\
x_4 & y_4 \\
x_3 & y_3 \\
x_2 & y_2 \\
x_1 & y_1 \\
\end{array}
\]
Incremental computation

\[ x^i = x_1 ... x_i \]
\[ \Phi(p, x \uparrow) \downarrow = y \text{ iff } \forall i < n : \Phi(p, x^i, n) = \downarrow y_{i+1} \]

\[
\begin{array}{ccc}
  x & = & y \\
  x_4 & = & y_4 \\
  x_3 & = & y_3 \\
  x_2 & = & y_2 \\
  x_1 & = & y_1 \\
\end{array}
\]
Incremental computation

\[ x^i = x_1 \ldots x_i \]

\[ \Phi(p, x \uparrow) \downarrow = y \text{ iff } \forall i < n : \Phi(p, x^i, n) = \downarrow y_{i+1} \]

\[
\begin{align*}
  x & \quad y \\
  = & \quad = \\
  x_4 & \quad y_4 \\
  x_3 & \quad y_3 \\
  x_2 & \quad y_2 \\
  x_1 & \quad y_1
\end{align*}
\]
Incremental computation

\[ x^i = x_1 \ldots x_i \]

\[ \Phi(p, x \uparrow) \downarrow = y \text{ iff } \forall i < n : \Phi(p, x^i, n) = \downarrow y_{i+1} \]

\[
\begin{align*}
  x & \quad y \\
  = & \quad = \\
  x_4 & \quad y_4 \\
  x_3 & \quad y_3 \\
  x_2 & \quad y_2 \\
  x_1 & \quad y_1
\end{align*}
\]
Incremental complexity

Definition

- **total conditional complexity (remind)**
  \[ K(x | y) = \min\{ l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \} \]

- **incremental conditional complexity**
  \[ K(x | y \uparrow) = \min\{ l(p) : \Phi(p, y \uparrow) \downarrow = x \}. \]

- **total incremental conditional complexity**
  \[ K(x \uparrow | y \uparrow) = \min\{ l(p) : \Phi(p, y \uparrow) \downarrow = x \land \forall z \in 2^n, \Phi(p, z \uparrow) \downarrow \} \]
Incremental complexity

**Definition**

- **total conditional complexity (remind)**
  \[ K(x|y) = \min \{ l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \} \]

- **incremental conditional complexity**
  \[ K(x|y \uparrow) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \} \]

- **total incremental conditional complexity**
  \[ K(x|y \uparrow) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \land \forall z \in 2^n, \Phi(p, z \uparrow) \downarrow \} \]
Incremental complexity

**Definition**

- **total conditional complexity (remind)**
  \[ K(x | y) = \min \{ l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \} \]

- **incremental conditional complexity**
  \[ K(x | y \uparrow) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \}. \]

- **total incremental conditional complexity**
  \[ K(x | y \uparrow) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \land \forall z \in 2^n, \Phi(p, z \uparrow) \downarrow \} \]
Incremental complexity

**Definition**

- **total conditional complexity (remind)**
  
  \[ K(x \mid y) = \min \{ l(p) : \Phi(p, y) \downarrow = x \land \forall z \in 2^n : \Phi(p, z) \downarrow \} \]

- **incremental conditional complexity**

  \[ K(x \mid y \uparrow) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \} \]

- **total incremental conditional complexity**

  \[ K(x \mid y \uparrow) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \land \forall z \in 2^n, \Phi(p, z \uparrow) \downarrow \} \]

Bruno Bauwens  Hypotheses and AIT
Proposition (Incremental coding bound)

If $P$ is a computable causal semimeasure, then

$$- \log P(x \mid y \uparrow) \geq K(x \mid y \uparrow).$$

Proposition (Incremental coding theorem)

$$K(x \mid y \uparrow) = \log m(x \mid y \uparrow) \pm O((\log k_{xy}))$$
Proposition (Incremental coding bound)

If $P$ is a computable causal semimeasure, then

$$- \log P(x|y \uparrow) \geq K(x|y \uparrow) + O((\log k_{xy}))$$

Proposition (Incremental coding theorem)

$$K(x|y \uparrow) = \log m(x|y \uparrow)$$
Well known decomposition

Theorem

\[ K(x, y) =^{+} K(x) + K(y|x^*) \]

with \( x^* = \text{shortest description for } x, \text{ witness of } K(x) \)
Incremental complexity: divergence

\[ K(x|y \uparrow^+) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \} \]
\[ K(x[y \uparrow^+) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \land \forall z \in 2^n : \Phi(p, z \uparrow) \downarrow \} \]

Trivial: \( x, y \) has constant bounded \( m \)-depth:

\[ K(x|y \uparrow) + K(y|x \uparrow^+) =^+ K(x, y) \]

Proposition

\[ \exists c > 0 \forall n \exists x, y \in 2^n : K(x|y \uparrow) + K(y|x \uparrow^+) - K(x, y) \geq^+ cn. \]
Incremental complexity: divergence

\[
K(x \mid y \uparrow^+) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \}
\]

\[
K(x \mid y \uparrow^+) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \land \forall z \in 2^n : \Phi(p, z \uparrow) \downarrow \}
\]

Trivial: $x, y$ has constant bounded $m$-depth:

\[
K(x \mid y \uparrow) + K(y \mid x \uparrow^+) =^+ K(x, y)
\]

**Proposition**

\[
\exists c > 0 \forall n \exists x, y \in 2^n : \ K(x \mid y \uparrow) + K(y \mid x \uparrow^+) - K(x, y) \geq^+ cn.
\]
Incremental\(^+\) complexity: divergence

\[
K(x|y \uparrow^+) = \min\{l(p) : \Phi(p, y \uparrow) \downarrow = x\}
\]
\[
K(x\upharpoonright y \uparrow^+) = \min\{l(p) : \Phi(p, y \uparrow) \downarrow = x \land \forall z \in 2^n : \Phi(p, z \uparrow) \downarrow\}
\]

Trivial: \(x, y\) has constant bounded \(m\)-depth:

\[
K(x|y \uparrow) + K(y|x \uparrow^+) =+ K(x, y)
\]

Proposition

\[
\exists c > 0 \forall n \exists x, y \in 2^n : K(x|y \uparrow) + K(y|x \uparrow^+) - K(x, y) \geq^+ cn.
\]
Incremental$^+$ complexity: divergence

\[
K(x|y \uparrow^+) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \}
\]
\[
K(x\uparrow | y \uparrow^+) = \min \{ l(p) : \Phi(p, y \uparrow) \downarrow = x \wedge \forall z \in 2^n : \Phi(p, z \uparrow) \downarrow \}
\]

Trivial: $x, y$ has constant bounded $m$-depth:

\[
K(x|y \uparrow) + K(y|x \uparrow^+) =^+ K(x, y)
\]

Proposition

\[
\exists c > 0 \forall n \exists x, y \in 2^n : \\
K(x|y \uparrow) + K(y|x \uparrow^+) - K(x, y) \geq^+ cn.
\]
Total conditional complexity: decomposition

Notation:

\[ K(y \uparrow x, p) = \min \{ l(q) : \Phi(q, x \uparrow, p) \downarrow y, \forall z \Phi(q, z \uparrow, p) \downarrow \} \]

Proposition

Let \( p \) be the minimal program in the definition of \( K(x \vert y) \):

\[ K(x \vert y) + K(y \vert p) \leq^+ K(x, y) + O(\log k_{xy}). \]

Let \( p \) be the minimal program in the definition of \( K(x \vert y \uparrow) \):

\[ K(x \vert y \uparrow) + K(y \vert x \uparrow, p) \leq^+ K(x, y) + O(\log k_{xy}). \]
Total conditional complexity: decomposition

Notation:

\[ K(y \uparrow x, p) = \min \{ l(q) : \Phi(q, x \uparrow p) \downarrow y, \forall z \Phi(q, z \uparrow p) \downarrow \} \]

Proposition

Let \( p \) be the minimal program in the definition of \( K(x \uparrow y) \):

\[ K(x \uparrow y) + K(y \mid p) \leq^+ K(x, y) + O(\log k_{xy}). \]

Let \( p \) be the minimal program in the definition of \( K(x \uparrow y) \):

\[ K(x \uparrow y) + K(y \uparrow x) \leq^+ K(x, y) + O(\log k_{xy}). \]
Total conditional complexity: decomposition

Notation:

\[ K(y|x \uparrow^+, p) = \min \{ l(q) : \Phi(q, x \uparrow^+, p) \downarrow y, \forall z \Phi(q, z \uparrow^+, p) \downarrow \} \]

Proposition

Let \( p \) be the minimal program in the definition of \( K(x|y) \):

\[ K(x|y) + K(y|p) \leq^+ K(x, y) + O(\log k_{xy}). \]

Let \( p \) be the minimal program in the definition of \( K(x|y \uparrow) \):

\[ K(x|y \uparrow) + K(y|x \uparrow^+, p) \leq^+ K(x, y) + O(\log k_{xy}). \]
Total conditional complexity: decomposition

Notation:

\[ K(y \uparrow^+, p) = \min\{ I(q) : \Phi(q, x \uparrow^+, p) \downarrow y, \forall z \Phi(q, z \uparrow^+, p) \downarrow \} \]

**Proposition**

Let \( p \) be the minimal program in the definition of \( K(x \uparrow y) \):

\[ K(x \uparrow y) + K(y \uparrow p) \leq^+ K(x, y) + O(\log k_{xy}). \]

Let \( p \) be the minimal program in the definition of \( K(x \uparrow y \uparrow) \):

\[ K(x \uparrow y \uparrow) + K(y \uparrow x \uparrow^+, p) \leq^+ K(x, y) + O(\log k_{xy}). \]
Question

\[ K(x|y \uparrow) + K(y|x \uparrow^+, p) =^+ K(x, y) + O(\log k_{xy}) \]

with \( p \) the minimal program in \( K(x|y \uparrow) \).
Generalization ?

Question

Let $S$ be an enumerable set of either finite or infinite, computable or enumerable sets in $\omega$. Let:

$$S_x = \arg \min \{K(S) : x \in S \in S\}$$

Does the following equation hold?

$$K(S_x) + K(x|S_x^*) = + K(x)$$

where $K(S) = I(S^*)$ and $S^*$ is either a minimal program that enumerates all elements of $S_x$ and halts, or a minimal program that enumerates all elements of $S_x$ and possibly continues computing.
Question

Let $S$ be an enumerable set of either finite or infinite, computable or enumerable sets in $\omega$. Let:

$$S_x = \arg\min\{K(S) : x \in S \in S\}$$

Does the following equation hold?

$$K(S_x) + K(x|S_x^*) =^+ K(x)$$

where $K(S) = l(S^*)$ and $S^*$ is either a minimal program that enumerates all elements of $S_x$ and halts, or a minimal program that enumerates all elements of $S_x$ and possibly continues computing.
Question

Let \( S \) be an enumerable set of either finite or infinite, computable or enumerable sets in \( \omega \). Let:

\[
S_x = \arg \min \{ K(S) : x \in S \in S \}
\]

Does the following equation hold?

\[
K(S_x) + K(x|S_x^*) =^+ K(x)
\]

where \( K(S) = l(S^*) \) and \( S^* \) is either a minimal program that enumerates all elements of \( S_x \) and halts, or a minimal program that enumerates all elements of \( S_x \) and possibly continues computing.
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
   - Ratio-test of universal semimeasures
   - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - Total conditional coding theorem
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
Algorithmic mutual information

\[
l(x; y) \equiv K(x) - K(x|y^*) \\
=^+ K(x) + K(y) - K(x, y) \\
=^+ I(y; x) \\
=^+ \log \frac{m(x, y)}{m(x)m(y)}
\]

Interpretation of \( I \) as ratio-test for independence (Levin).
Time series

'Origin' of $K(x, y)$ from three sources.

$I(x; y)$ as a sum of (the three arrows in the middle):
- Information flow from $x$ to $y$.
- Information flow from $y$ to $x$.
- Information from a common source.
Information transfer and instantaneous common information

\[ IT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]
\[ TIT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]
\[ IT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]
\[ TIT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]
\[ ITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]
\[ TITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]

with \( p \) the withness of \( K(x|y \uparrow) \) and \( K(x|y \uparrow) \).
Information transfer and instantaneous common information

**total information transfer**

**total conditional instantaneous common information**

\[ IT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]

\[ TIT(x \leftarrow y) = K(x) - K(x[y \uparrow]) \]

\[ IT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]

\[ TIT(x \uparrow; y \uparrow) = K(x[y \uparrow]) - K(x[y \uparrow^+]) \]

\[ ITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]

\[ TITc(x \uparrow; y \uparrow) = K(x[y \uparrow]) - K(x[y \uparrow^+, p]) \]

with \( p \) the withness of \( K(x[y \uparrow]) \)
Information transfer and instantaneous common information

Total information transfer
Total conditional instantaneous common information

\[ IT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]
\[ TIT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]
\[ IT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]
\[ TIT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]
\[ ITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]
\[ TITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]

With \( p \) the withness of \( K(x|y \uparrow) \) and \( K(x|y \uparrow) \).
Information transfer and instantaneous common information

**Total information transfer**

\[ IT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]

**Total conditional information**

\[ TIT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]

**Instantaneous common information**

\[ IT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]

**Total information transfer**

\[ TIT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]

\[ ITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]

\[ TITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]

with \( p \) the withness of \( K(x|y \uparrow) \) and \( K(x|y \uparrow^+) \).
Information transfer and instantaneous common information

- **Total information transfer**
  - \( I_T(x \leftarrow y) = K(x) - K(x|y \uparrow) \)
  - \( TIT(x \leftarrow y) = K(x) - K(x|y \uparrow) \)
  - \( I_T(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \)
  - \( TIT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \)

- **Total conditional instantaneous common information**
  - \( I_{Tc}(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \)
  - \( TITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \)

with \( p \) the withness of \( K(x|y \uparrow) \) and \( K(x|y \uparrow) \).
Information transfer and instantaneous common information

Total information transfer

Total conditional instantaneous common information

\[ IT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]

\[ TIT(x \leftarrow y) = K(x) - K(x|y \uparrow) \]

\[ IT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]

\[ TIT(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+) \]

\[ ITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]

\[ TITc(x \uparrow; y \uparrow) = K(x|y \uparrow) - K(x|y \uparrow^+, p) \]

with \( p \) the withness of \( K(x|y \uparrow) K(x|y \uparrow) \)
Corollary

\[ TIT(y \leftarrow x) + TIT(x \leftarrow y) + TITc(x \uparrow; y \uparrow) = I(x; y) \pm O(\log k_{xy}) \]

\[ TITc(x; y) = TITc(y; x) \pm O(\log k_{xy}). \]

Question

How symmetric is ITc(x \uparrow; y \uparrow)?

For multi-symbol tapes:

\[ I(x \uparrow; y \uparrow) - I(y \uparrow; x \uparrow) \geq o(\log n). \]
Corollary

\[ TIT(y \leftarrow x) + TIT(x \leftarrow y) + TITc(x \uparrow; y \uparrow) \]
\[ = I(x; y) \pm O(\log k_{xy}) \]

\[ TITc(x; y) = TITc(y; x) \pm O(\log k_{xy}). \]

Question

*How symmetric is \( ITc(x \uparrow; y \uparrow) \)?*

For multi-symbol tapes:

\[ I(x \uparrow; y \uparrow) - I(y \uparrow; x \uparrow) \geq o(\log n). \]
Corollary

\[
TIT(y \leftarrow x) + TIT(x \leftarrow y) + TITc(x \uparrow; y \uparrow) \\
=^+ I(x; y) \pm O(\log k_{xy}) \\
TITc(x; y) = TITc(y; x) \pm O(\log k_{xy}).
\]

Question

*How symmetric is \( ITc(x \uparrow; y \uparrow) \)?*

For multi-symbol tapes:

\[
I(x \uparrow; y \uparrow) - I(y \uparrow; x \uparrow) \geq o(\log n).
\]
Corollary

\[
TIT(y \leftarrow x) + TIT(x \leftarrow y) + TITc(x \uparrow; y \uparrow) \\
=+ I(x; y) \pm O(\log k_{xy}) \\
TITc(x; y) = TITc(y; x) \pm O(\log k_{xy}).
\]

Question

*How symmetric is ITc(x \uparrow; y \uparrow) ?*

For multi-symbol tapes:

\[
l(x \uparrow; y \uparrow) - l(y \uparrow; x \uparrow) \geq o(\log n).
\]
**Corollary**

\[
TIT(y \leftarrow x) + TIT(x \leftarrow y) + TITc(x \uparrow; y \uparrow) \\
=^+ I(x; y) \pm O(\log k_{xy}) \\
TITc(x; y) = TITc(y; x) \pm O(\log k_{xy}).
\]

**Question**

*How symmetric is ITc(x \uparrow; y \uparrow)?*

For multi-symbol tapes:

\[
I(x \uparrow; y \uparrow) - I(y \uparrow; x \uparrow) \geq o(\log n).
\]
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
   - Ratio-test of universal semimeasures
   - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - Total conditional coding theorem
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
Let $P$ be a semimeasure over $\omega$.

$$d : \omega \cup \{-1\} \rightarrow \omega \cup \{-1\} \text{ is a } P\text{-sumtest iff}$$

$$\sum_{x \in \omega} P(x)2^{d(x)} \leq 1$$

- Identity testing: “Is $x$ typical for $P$?”
- Largest $d$ in some computability class?
Let $P$ be a semimeasure over $\omega$.

$d : \omega \cup \{-1\} \to \omega \cup \{-1\}$ is a $P$-sumtest iff

$$\sum_{x \in \omega} P(x)2^{d(x)} \leq 1$$

- Identity testing: “Is $x$ typical for $P$?”
- Largest $d$ in some computability class?
Let $P$ be a semimeasure over $\omega$.

$d : \omega \cup \{-1\} \rightarrow \omega \cup \{-1\}$ is a $P$-sumtest iff

$$\sum_{x \in \omega} P(x)2^{d(x)} \leq 1$$

- Identity testing: “Is $x$ typical for $P$?”
- Largest $d$ in some computability class?
Let $P$ be a semimeasure over $\omega$.

$d : \omega \cup \{-1\} \rightarrow \omega \cup \{-1\}$ is a $P$-sumtest iff

$$\sum_{x \in \omega} P(x)2^{d(x)} \leq 1$$

- Identity testing: “Is $x$ typical for $P$?”
- Largest $d$ in some computability class?
A function $f$ dominates $g$ ($f \geq^+ g$) iff

$$\exists c \forall x : f(x) + c \geq g(x).$$

A function $f$ is universal in a set $S$ iff $f \in S$ and

$$\forall g \in S : f \geq^+ g.$$
Universality

- A function $f$ dominates $g$ ($f \geq^+ g$) iff
  \[ \exists c \forall x : f(x) + c \geq g(x). \]

- A function $f$ is universal in a set $S$ iff $f \in S$ and
  \[ \forall g \in S : f \geq^+ g. \]
Universality

- A function $f$ dominates $g$ ($f \geq^+ g$) iff
  \[ \exists c \forall x : f(x) + c \geq g(x). \]

- A function $f$ is universal in a set $S$ iff $f \in S$ and
  \[ \forall g \in S : f \geq^+ g. \]
Universality theorem

Proposition

Let $d$ be an $m$-sumtest,

- if $d$ is computable or enumerable then

  $$d(x) \leq^+ 2K(d),$$

- if $d$ is co-enumerable then

  $$d(x) \leq^+ \log l(x) + 4 \log \log l(x),$$

and

$$\exists d' \forall n \exists x, y \in \{0, 1\}^n : d'(x, y) - d(x, y) \geq^+ \log l(x) - O(\log^1 l(x)).$$
Proposition

Let $d$ be an $m$-sumtest,

- if $d$ is computable or enumerable than
  \[ d(x) \leq + 2K(d), \]

- if $d$ is co-enumerable than
  \[ d(x) \leq + \log l(x) + 4 \log \log l(x), \]

and

- \[ \exists d' \forall n \exists x, y \in 1^n : d'(x, y) - d(x, y) \geq + \log l(x) - O(\log^1 l(x)). \]
University theorem

**Proposition**

Let \( d \) be an \( m \)-sumtest,

- if \( d \) is computable or enumerable then
  \[
  d(x) \leq 2K(d),
  \]

- if \( d \) is co-enumerable then
  \[
  d(x) \leq \log l(x) + 4 \log \log l(x),
  \]

and

\[
\exists d' \forall n \exists x, y \in 1^n : d'(x, y) - d(x, y) \geq \log l(x) - O(\log^1 l(x)).
\]
Universality theorem

**Proposition**

*Let $d$ be an $m$-sumtest,*

- *if $d$ is computable or enumerable than*

  $$d(x) \leq 2K(d),$$

- *if $d$ is co-enumerable than*

  $$d(x) \leq \log l(x) + 4 \log \log l(x),$$

  *and*

  $$\exists d' \forall n \exists x, y \in 1^n : d'(x, y) - d(x, y) \geq \log l(x) - O(\log^1 l(x)).$$
Given: co-enumerable $m$-sumtest $d$

Required: $x$ and $d'$ such that $d(x) - d'(x) \geq ...$

- Few $x$ have high $d_0(x)$ (sumtest).
- Choose $x$ such that $m_t(x)$ remains constant and increases at large $t = s$: $K^+(s) \gg K(d)$.
- $d_0$ can impossibly distinguish $x$ in a set of $y$ with constant $m_t(y)$. Therefore, $d_0(x)$ is low.
- Construct $d'$ that knows $s$ and therefore can reserve a high $d'_0(x)$
- ... iterate ...

Bruno Bauwens  Hypotheses and AIT
Subgoals of proof

**Given:** co-enumerable $m$-sumtest $d$

**Required:** $x$ and $d'$ such that $d(x) - d'(x) \geq ...$

- Few $x$ have high $d_0(x)$ (sumtest).
- Choose $x$ such that $m_t(x)$ remains constant and increases at large $t = s$: $K^+(s) \gg K(d)$.
- $d_0$ can impossibly distinguish $x$ in a set of $y$ with constant $m_t(y)$. Therefore, $d_0(x)$ is low.
- Construct $d'$ that knows $s$ and therefore can reserve a high $d_0'(x)$
- ... iterate ...
Subgoals of proof

**Given:** co-enumerable $m$-sumtest $d$

**Required:** $x$ and $d'$ such that $d(x) - d'(x) \geq \ldots$

- Few $x$ have high $d_0(x)$ (sumtest).
- Choose $x$ such that $m_t(x)$ remains constant and increases at large $t = s$: $K^+(s) \gg K(d)$.
- $d_0$ can impossibly distinguish $x$ in a set of $y$ with constant $m_t(y)$. Therefore, $d_0(x)$ is low.
- Construct $d'$ that knows $s$ and therefore can reserve a high $d'_0(x)$
- ... iterate ...
Subgoals of proof

**Given:** co-enumerable $m$-sumtest $d$

**Required:** $x$ and $d'$ such that $d(x) - d'(x) \geq \ldots$

- Few $x$ have high $d_0(x)$ (sumtest).
- Choose $x$ such that $m_t(x)$ remains constant and increases at large $t = s$: $K^+(s) \gg K(d)$.
- $d_0$ can impossibly distinguish $x$ in a set of $y$ with constant $m_t(y)$. Therefore, $d_0(x)$ is low.
- Construct $d'$ that knows $s$ and therefore can reserve a high $d'_0(x)$
- ... iterate ...
Subgoals of proof

**Given:** co-enumerable $m$-sumtest $d$

**Required:** $x$ and $d'$ such that $d(x) - d'(x) \geq \ldots$

- Few $x$ have high $d_0(x)$ (sumtest).
- Choose $x$ such that $m_t(x)$ remains constant and increases at large $t = s$: $K^+(s) \gg K(d)$.
- $d_0$ can impossibly distinguish $x$ in a set of $y$ with constant $m_t(y)$. Therefore, $d_0(x)$ is low.
- Construct $d'$ that knows $s$ and therefore can reserve a high $d'_0(x)$
- ... iterate ...
Subgoals of proof

**Given:** co-enumerable $m$-sumtest $d$

**Required:** $x$ and $d'$ such that $d(x) - d'(x) \geq \ldots$

- Few $x$ have high $d_0(x)$ (sumtest).
- Choose $x$ such that $m_t(x)$ remains constant and increases at large $t = s$: $K^+(s) \gg K(d)$.
- $d_0$ can impossibly distinguish $x$ in a set of $y$ with constant $m_t(y)$. Therefore, $d_0(x)$ is low.
- Construct $d'$ that knows $s$ and therefore can reserve a high $d'_0(x)$
- ... iterate ...
Subgoals of proof

**Given:** co-enumerable $m$-sumtest $d$

**Required:** $x$ and $d'$ such that $d(x) - d'(x) \geq ...$

- Few $x$ have high $d_0(x)$ (sumtest).
- Choose $x$ such that $m_t(x)$ remains constant and increases at large $t = s$: $K^+(s) \gg K(d)$.
- $d_0$ can impossibly distinguish $x$ in a set of $y$ with constant $m_t(y)$. Therefore, $d_0(x)$ is low.
- Construct $d'$ that knows $s$ and therefore can reserve a high $d'_0(x)$

... iterate ...
Subgoals of proof

**Given:** co-enumerable \( m \)-sumtest \( d \)

**Required:** \( x \) and \( d' \) such that \( d(x) - d'(x) \geq \ldots \)

- Few \( x \) have high \( d_0(x) \) (sumtest).
- Choose \( x \) such that \( m_t(x) \) remains constant and increases at large \( t = s \): \( K^+(s) \gg K(d) \).
- \( d_0 \) can impossibly distinguish \( x \) in a set of \( y \) with constant \( m_t(y) \). Therefore, \( d_0(x) \) is low.
- Construct \( d' \) that knows \( s \) and therefore can reserve a high \( d'_0(x) \)
- \( \ldots \) iterate \( \ldots \)
Gradual compressible strings

\[ K_s(x) \]

\[ 1_{a-1} \]
\[ 1_a \]
\[ 1_{a+1} \]

\[ S_{a-1} \quad T(S_{a-1}) \]
\[ S_a \quad T(S_a) \]
Gradual compressible strings

$K_s(x)$
Constructed $x$ has high minimal sufficient statistics. Any co-enumerable test bounded by minimal sufficient statistic within logarithmic term.
Outline

1. Composite Hypothesis tests
   - Motivation
   - Procedure
   - Ratio-test of universal semimeasures
   - Criterion for universality

2. Influence tests
   - Causal semimeasures
   - Total conditional coding theorem
   - Incremental coding theorem

3. Decompositions
   - Decompositions of algorithmic complexity
   - Decompositions of mutual information

4. Randomness tests in hypotheses
   - Sumtests
   - Independence tests
A $(P, Q)$-independence test $d$ is a function $d : \omega \cup \{-1\}^2 \rightarrow \omega \cup \{-1\}$ that satisfies the following condition:

$$\sum_{x, y \in \omega} P(x) Q(y) 2^{d(x, y)} \leq 1.$$
Independence test

Definition

\[ d : \omega \cup \{-1\}^2 \rightarrow \omega \cup \{-1\} \text{ is a } (P, Q)\text{-independence test iff} \]

\[ \sum_{x, y \in \omega} P(x)Q(y)2^{d(x, y)} \leq 1. \]
Proposition

Let \( d \) be an \((m, m)\)-independence test,

- if \( d \) is computable or enumerable, then
  \[
d(x, y) \leq 2K(d)
  \]
- if \( d \) is co-enumerable,

  \[
  \exists d' \forall n \exists x, y \in 2^n : d'(x, y) - d(x, y) \geq \log l(x) - O(\log^2 l(x)),
  \]
  \[
  d'(x, y) - d(x, y) \geq l(x) - O(\log l(x)).
  \]
Summary

- Influence: Ideal limit point for improving sequence of algorithms exist
- Independence: Ideal limit point does not exist
- Decomposition for Kolmogorov complexity and mutual information
- Answering simple questions in statistics, often involves the use of 'Halting information’ and ’notions of computational depth'. 
Summary

- **Influence**: Ideal limit point for improving sequence of algorithms exist
- **Independence**: Ideal limit point does not exist
- **Decomposition for Kolmogorov complexity and mutual information**
- Answering simple questions in statistics, often involves the use of 'Halting information' and 'notions of computational depth'.
Summary

- **Influence:** Ideal limit point for improving sequence of algorithms exist
- **Independence:** Ideal limit point does not exist
- **Decomposition:** for Kolmogorov complexity and mutual information
- **Answering simple questions in statistics,** often involves the use of ‘Halting information’ and ‘notions of computational depth’.