Unlabeled Compression Schemes

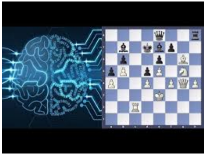
Dömötör Pálvölgyi

MTA-ELTE CoGe, Budapest

28 June 2018

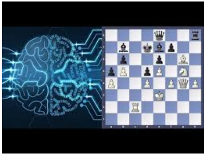
Learning theory

Learning theory



AlphaZero

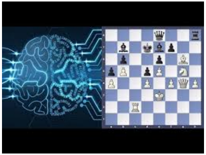
Learning theory



AlphaZero

AlphaGo

Learning theory

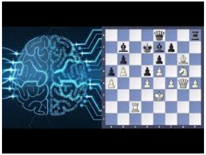


AlphaZero

AlphaGo

Self-driving cars

Learning theory



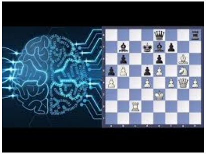
AlphaZero

Self-driving cars

AlphaGo

Powered by Neural Networks.

Learning theory



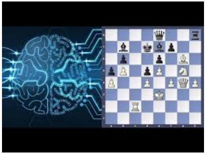
AlphaZero

Self-driving cars

Powered by Neural Networks.

Key to compress data

Learning theory



AlphaZero

Self-driving cars

Powered by Neural Networks.

Key to compress data — experts guess lookup table (oracle) for chess compressable to ≈300MB.

Compress this

can speak russian

mathematician

has not been
on the moon

male

participates in
this conference

family name starts with K

Compress this

can speak russian

mathematician

has not been on the moon

male

---

below this line implies everything above line

participates in this conference

family name starts with K

Compressing a sample

Compressing a sample

Suppose we want to learn an interval from some sample.

Compressing a sample

Suppose we want to learn an interval from some sample.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Decompression for non-sample points can change.

Compressing a sample

Suppose we want to learn an interval from some sample.

We can throw away unimportant information.

What counts as unimportant?

Can throw away more if goal is to be able to recover sample points.

Decompression for non-sample points can change.

Sample compression scheme

Sample compression scheme

S: a $\{+,-\}$ or $\{+,-,.\}$ sequence on elements of base set.

Sample compression scheme

S: a $\{+,-\}$ or $\{+,-,.\}$ sequence on elements of base set.

T is a mask of S if obtained from S by replacing any number of digits by (the same number of) .'s.

Sample compression scheme

S: a $\{+,-\}$ or $\{+,-,.\}$ sequence on elements of base set.

T is a mask of S if obtained from S by replacing any number of digits by (the same number of) .'s.

Concept	*Sample*	*Compressed sample*	*Decoding*
$F \in \mathcal{F}$	S	$\alpha(S)$	$\beta(\alpha(S))$
$\pm_F$	mask of F	mask of S	$\notin \mathcal{F}$ allowed
$\{+,-\}$'s	$\{+,-,.\}$'s	$\{+,-,.\}$'s	$\{+,-\}$'s

Sample compression scheme

S: a $\{+,-\}$ or $\{+,-,.\}$ sequence on elements of base set.

T is a mask of S if obtained from S by replacing any number of digits by (the same number of) .'s.

Concept	*Sample*	*Compressed sample*	*Decoding*
$F \in \mathcal{F}$	S	$\alpha(S)$	$\beta(\alpha(S))$
$\pm_F$	mask of F	mask of S	$\notin \mathcal{F}$ allowed
$\{+,-\}$'s	$\{+,-,.\}$'s	$\{+,-,.\}$'s	$\{+,-\}$'s

Compression scheme is valid if $\beta(\alpha(S))$ matches S wherever $S \neq .$

Sample compression scheme

Given $\mathcal{F}$ set system (aka. concept class), find sample compression scheme such that value at any sample point can be recovered.

Example. $\mathcal{F} = \binom{\mathbb{N}}{\leq d}$,

Sample compression scheme

Given $\mathcal{F}$ set system (aka. concept class), find sample compression scheme such that value at any sample point can be recovered.

Example. $\mathcal{F} = \binom{\mathbb{N}}{\leq d}$,

Sample: $\pm_S$ with at most d $+$'s.

Sample compression scheme

Given $\mathcal{F}$ set system (aka. concept class), find sample compression scheme such that value at any sample point can be recovered.

Example. $\mathcal{F} = \binom{\mathbb{N}}{\leq d}$,

Sample: $\pm_S$ with at most d $+$'s.

Compression: keep only $+$'s (rest are .).

Sample compression scheme

Given $\mathcal{F}$ set system (aka. concept class), find sample compression scheme such that value at any sample point can be recovered.

Example. $\mathcal{F} = \binom{\mathbb{N}}{\leq d}$,

Sample: $\pm_S$ with at most d $+$'s.

Compression: keep only $+$'s (rest are .).

This way everything compresses to $\leq d$ sample points.

Sample compression scheme

Given $\mathcal{F}$ set system (aka. concept class), find sample compression scheme such that value at any sample point can be recovered.

Example. $\mathcal{F} = \binom{\mathbb{N}}{\leq d}$,

Sample: $\pm_S$ with at most d $+$'s.

Compression: keep only $+$'s (rest are .).

This way everything compresses to $\leq d$ sample points.

Notice that here (and also for intervals) the $\pm$ labeling of compressed sample can be ignored in decompression.

Sample compression scheme

Given $\mathcal{F}$ set system (aka. concept class), find sample compression scheme such that value at any sample point can be recovered.

Example. $\mathcal{F} = \binom{\mathbb{N}}{\leq d}$,

Sample: $\pm_S$ with at most d $+$'s.

Compression: keep only $+$'s (rest are .).

This way everything compresses to $\leq d$ sample points.

Notice that here (and also for intervals) the $\pm$ labeling of compressed sample can be ignored in decompression.

Sample compression scheme

Given $\mathcal{F}$ set system (aka. concept class), find sample compression scheme such that value at any sample point can be recovered.

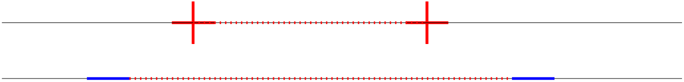
Example. $\mathcal{F} = \binom{\mathbb{N}}{\leq d}$,

Sample: $\pm_S$ with at most d $+$'s.

Compression: keep only $+$'s (rest are .).

This way everything compresses to $\leq d$ sample points.

Notice that here (and also for intervals) the $\pm$ labeling of compressed sample can be ignored in decompression.



Definitions: VC, LCS, UCS

Definitions: VC, LCS, UCS

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \; \exists F \in \mathcal{F} \; X \cap F = Y\}.$$

Definitions: VC, LCS, UCS

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \ \exists F \in \mathcal{F} \ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

Definitions: $\mathrm{VC}, \mathrm{LCS}, \mathrm{UCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

Definitions: VC, LCS, UCS

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

By definition $\mathrm{LCS} \leq \mathrm{UCS}$.

$\mathrm{VC} \leq \mathrm{UCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \; \exists F \in \mathcal{F} \; X \cap F = Y\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

Claim. By pigeonhole principle $\mathrm{VC} \leq \mathrm{UCS}$.

$\mathrm{VC} \leq \mathrm{UCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \; \exists F \in \mathcal{F} \; X \cap F = Y\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

Claim. By pigeonhole principle $\mathrm{VC} \leq \mathrm{UCS}$.

Proof. Take $|S| = d$.

VC $\leq$ UCS

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

Claim. By pigeonhole principle $\mathrm{VC} \leq \mathrm{UCS}$.

Proof. Take $|S| = d$.
There are 2^d labelings of S and 2^d (unlabeled) subsets of S. □

VC $\not\leq$ LCS

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \ \exists F \in \mathcal{F} \ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

For $\mathcal{F} = \{+,-\}^2$, $\mathrm{VC} = 2$, $\mathrm{LCS} = 1$:

$\mathrm{VC} \not\leq \mathrm{LCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme

For $\mathcal{F} = \{+,-\}^2$, $\mathrm{VC} = 2$, $\mathrm{LCS} = 1$:

Sample	*Compression*	*Decoding*
$\{+,+\}$	$\{+,.\}$	$\{+,+\}$
$\{-,-\}$	$\{-,.\}$	$\{-,-\}$
$\{+,-\}$	$\{.,-\}$	$\{+,-\}$
$\{-,+\}$	$\{.,+\}$	$\{-,+\}$
$\{+,.\}$	$\{+,.\}$	$\{+,+\}$
$\{.,-\}$	$\{.,-\}$	$\{+,-\}$
$\emptyset$	$\emptyset$	any

$\mathrm{VC} \leq 5 \cdot \mathrm{LCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \ \exists F \in \mathcal{F} \ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

Claim (Littlestone-Warmuth). $\mathrm{VC} \leq 5 \cdot \mathrm{LCS}$.

$\mathrm{VC} \leq 5 \cdot \mathrm{LCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

Claim (Littlestone-Warmuth). $\mathrm{VC} \leq 5 \cdot \mathrm{LCS}$.

Proof. Take $|S| = \mathrm{VC} = d$.

$\mathrm{VC} \leq 5 \cdot \mathrm{LCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

Claim (Littlestone-Warmuth). $\mathrm{VC} \leq 5 \cdot \mathrm{LCS}$.

Proof. Take $|S| = \mathrm{VC} = d$.

2^d labelings of S, and labeled subsets of S of size $\leq d/5 =$

$$\sum_{i=0}^{d/5} \binom{d}{i} 2^i \leq \left(\frac{ed}{d/5}\right)^{d/5} 2^{d/5} \leq (10e)^{d/5} < 2^d. \quad \square$$

Summary: $\mathrm{VC}, \mathrm{LCS}, \mathrm{UCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \ \exists F \in \mathcal{F} \ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$$\mathrm{VC}/5 \leq \mathrm{LCS} \leq \mathrm{UCS}, \mathrm{VC} \leq \mathrm{UCS}$$

Summary: $\mathrm{VC}, \mathrm{LCS}, \mathrm{UCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$$\mathrm{VC}/5 \leq \mathrm{LCS} \leq \mathrm{UCS}, \mathrm{VC} \leq \mathrm{UCS}$$

Conjecture (Littlestone-Warmuth). $\mathrm{LCS} = O(\mathrm{VC})$

Summary: $\mathrm{VC}, \mathrm{LCS}, \mathrm{UCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \ \exists F \in \mathcal{F} \ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$$\mathrm{VC}/5 \leq \mathrm{LCS} \leq \mathrm{UCS}, \mathrm{VC} \leq \mathrm{UCS}$$

Conjecture (Littlestone-Warmuth). $\mathrm{LCS} = O(\mathrm{VC})$ — worth 600$!

Summary: VC, LCS, UCS

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$$\mathrm{VC}/5 \leq \mathrm{LCS} \leq \mathrm{UCS}, \mathrm{VC} \leq \mathrm{UCS}$$

Conjecture (Littlestone-Warmuth). $\mathrm{LCS} = O(\mathrm{VC})$ — worth 600$!

It is even possible that $\mathrm{LCS} \leq \mathrm{VC}$.

Summary: $\mathrm{VC}, \mathrm{LCS}, \mathrm{UCS}$

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

$\mathrm{LCS}(\mathcal{F})$: the size of the smallest labeled compression scheme, i.e.,

$$\min_{\alpha \text{ valid compression scheme for } \mathcal{F}} \max\{|\alpha(S^{\pm})|\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$$\mathrm{VC}/5 \leq \mathrm{LCS} \leq \mathrm{UCS}, \mathrm{VC} \leq \mathrm{UCS}$$

Conjecture (Littlestone-Warmuth). $\mathrm{LCS} = O(\mathrm{VC})$ — worth 600$!

It is even possible that $\mathrm{LCS} \leq \mathrm{VC}$. Maybe $\mathrm{UCS} = \mathrm{VC}$? (Floyd-W)

UCS $\leq$ VC?

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \; \exists F \in \mathcal{F} \; X \cap F = Y\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$\mathrm{UCS} \leq \mathrm{VC} = d$ true for

$\mathrm{UCS} \leq \mathrm{VC}$?

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \; \exists F \in \mathcal{F} \; X \cap F = Y\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$\mathrm{UCS} \leq \mathrm{VC} = d$ true for $d = 1$,

UCS $\leq$ VC?

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \; \exists F \in \mathcal{F} \; X \cap F = Y\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$\mathrm{UCS} \leq \mathrm{VC} = d$ true for $d = 1$, and for $|\mathcal{F}| = \sum_{i=0}^{d} \binom{d}{i}$.

UCS $\leq$ VC?

VC($\mathcal{F}$): VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X\ \exists F \in \mathcal{F}\ X \cap F = Y\}.$$

UCS($\mathcal{F}$): the size of the smallest unlabeled compression scheme.

$\mathrm{UCS} \leq \mathrm{VC} = d$ true for $d = 1$, and for $|\mathcal{F}| = \sum_{i=0}^{d} \binom{d}{i}$.

Attack on LW/FW-conjecture: Can we embed any $\mathcal{F}$ into a maximum $\mathcal{F}$' such that $\mathrm{VC}(\mathcal{F}') \leq O(\mathrm{VC}(\mathcal{F}))$?

UCS $\leq$ VC?

$\mathrm{VC}(\mathcal{F})$: VC-dimension of $\mathcal{F}$, i.e.,

$$\max\{|X| \mid \forall Y \subset X \ \exists F \in \mathcal{F} \ X \cap F = Y\}.$$

$\mathrm{UCS}(\mathcal{F})$: the size of the smallest unlabeled compression scheme.

$\mathrm{UCS} \leq \mathrm{VC} = d$ true for $d = 1$, and for $|\mathcal{F}| = \sum_{i=0}^{d} \binom{d}{i}$.

Attack on LW/FW-conjecture: Can we embed any $\mathcal{F}$ into a maximum $\mathcal{F}$' such that $\mathrm{VC}(\mathcal{F}') \leq O(\mathrm{VC}(\mathcal{F}))$? — OPEN.

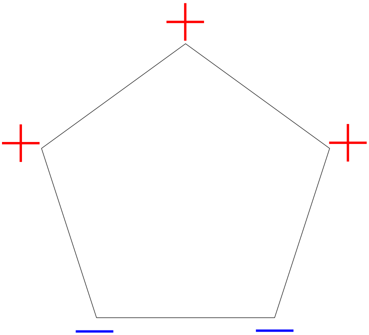
UCS $\not\leq$ VC

Example (P-Tardos). For $\mathcal{F} = C_5$, $\mathrm{VC} = 2$ but $\mathrm{UCS} = 3$.

UCS $\not\leq$ VC

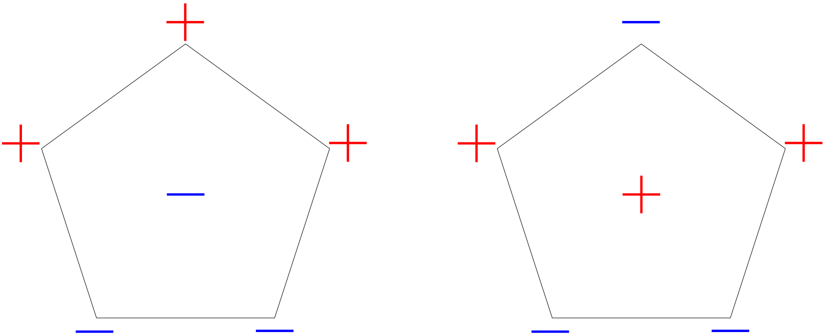
Example (P-Tardos). For $\mathcal{F} = C_5$, $\mathrm{VC} = 2$ but $\mathrm{UCS} = 3$.

$|C_5| = 10$, the 5 rotations of the following sets:



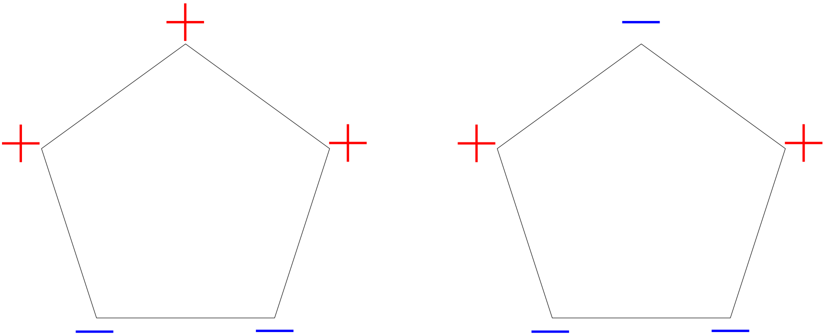
UCS $\not\leq$ VC

Family C_5 can be made more symmetric by extending onto one more element (still $\mathrm{VC} = 2$, $\mathrm{UCS} = 3$, $|W_6| = 10$), 5 rotations of:



UCS $\not\leq$ VC

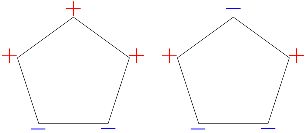
Example (P-Tardos). For $\mathcal{F} = C_5$, $\mathrm{VC} = 2$ but $\mathrm{UCS} = 3$.
$|C_5| = 10$, the 5 rotations of the following sets:



$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.

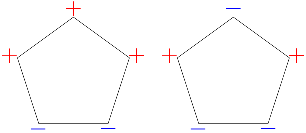
$|C_5| = 10$, the 5 rotations of the following sets:



$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.
There are 7 restriction of C_5 to any triple.

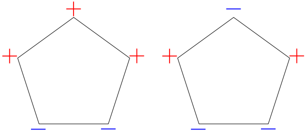
$|C_5| = 10$, the 5 rotations of the following sets:



$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.
There are 7 restriction of C_5 to any triple.
All 7 compressions must give one of these.

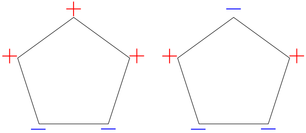
$|C_5| = 10$, the 5 rotations of the following sets:



$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.
There are 7 restriction of C_5 to any triple.
All 7 compressions must give one of these.
Any $|X| \leq 3$ set is decoded to such a valid triple $Y^{\pm}$ on each $Y \supset X$ with $|Y| = 3$.

$|C_5| = 10$, the 5 rotations of the following sets:

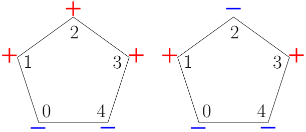


$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.

There are 7 restriction of C_5 to any triple.
All 7 compressions must give one of these.
Any $|X| \leq 3$ set is decoded to such a valid triple $Y^{\pm}$ on each $Y \supset X$ with $|Y| = 3$.
E.g., decoding of $X = \{2\}$ cannot give $\{1^-, 2^-, 3^-\}$, $\{0^+, 2^+, 4^+\}$ etc.

$|C_5| = 10$, the 5 rotations of the following sets:



$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.

There are 7 restriction of C_5 to any triple.

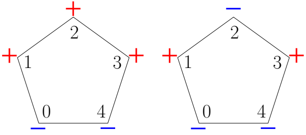
All 7 compressions must give one of these.

Any $|X| \le 3$ set is decoded to such a valid triple $Y^{\pm}$ on each $Y \supset X$ with $|Y| = 3$.

E.g., decoding of $X = \{2\}$ cannot give $\{1^-, 2^-, 3^-\}$, $\{0^+, 2^+, 4^+\}$ etc.

It also cannot be $\{1^-, 2^+, 3^-\}$ because some $X \subset \{1, 3\}$ has to give $\{1^-, 3^-\}$.

$|C_5| = 10$, the 5 rotations of the following sets:



$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.

There are 7 restriction of C_5 to any triple.

All 7 compressions must give one of these.

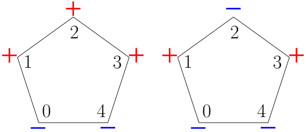
Any $|X| \leq 3$ set is decoded to such a valid triple $Y^{\pm}$ on each $Y \supset X$ with $|Y| = 3$.

E.g., decoding of $X = \{2\}$ cannot give $\{1^-, 2^-, 3^-\}$, $\{0^+, 2^+, 4^+\}$ etc.

It also cannot be $\{1^-, 2^+, 3^-\}$ because some $X \subset \{1, 3\}$ has to give $\{1^-, 3^-\}$.

Thus $\{2\}$ is decoded to $\{0^-, 1^+, 2^?, 3^+, 4^-\}$.

$|C_5| = 10$, the 5 rotations of the following sets:

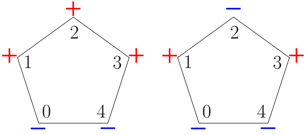


$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.

There are 7 restriction of C_5 to any triple.
All 7 compressions must give one of these.
Any $|X| \leq 3$ set is decoded to such a valid triple $Y^{\pm}$ on each $Y \supset X$ with $|Y| = 3$.
E.g., decoding of $X = \{2\}$ cannot give $\{1^-, 2^-, 3^-\}$, $\{0^+, 2^+, 4^+\}$ etc.
It also cannot be $\{1^-, 2^+, 3^-\}$ because some $X \subset \{1, 3\}$ has to give $\{1^-, 3^-\}$.
Thus $\{2\}$ is decoded to $\{0^-, 1^+, 2^?, 3^+, 4^-\}$.
If $\{2\}$ is decoded to $\{0^-, 1^+, 2^+, 3^+, 4^-\}$ and $\{1\}$ is decoded to $\{0^+, 1^+, 2^+, 3^-, 4^-\}$, then they clash at $\{1, 2, 4\}$.

$|C_5| = 10$, the 5 rotations of the following sets:

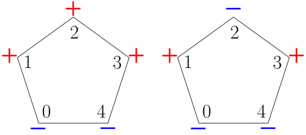


$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.

There are 7 restriction of C_5 to any triple.
All 7 compressions must give one of these.
Any $|X| \leq 3$ set is decoded to such a valid triple $Y^{\pm}$ on each $Y \supset X$ with $|Y| = 3$.
E.g., decoding of $X = \{2\}$ cannot give $\{1^-, 2^-, 3^-\}$, $\{0^+, 2^+, 4^+\}$ etc.
It also cannot be $\{1^-, 2^+, 3^-\}$ because some $X \subset \{1, 3\}$ has to give $\{1^-, 3^-\}$.
Thus $\{2\}$ is decoded to $\{0^-, 1^+, 2^?, 3^+, 4^-\}$.
If $\{2\}$ is decoded to $\{0^-, 1^+, 2^+, 3^+, 4^-\}$ and $\{1\}$ is decoded to $\{0^+, 1^+, 2^+, 3^-, 4^-\}$, then they clash at $\{1, 2, 4\}$.
If $\{2\}$ is decoded to $\{0^-, 1^+, 2^-, 3^+, 4^-\}$ and $\{0\}$ is decoded to $\{0^-, 1^+, 2^-, 3^-, 4^+\}$, then they clash at $\{0, 1, 2\}$.

$|C_5| = 10$, the 5 rotations of the following sets:

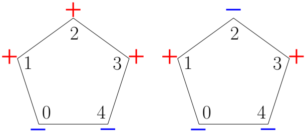


$\mathrm{UCS}(C_5) = 3$

Claim. $\mathrm{UCS}(C_5) = 3$.

There are 7 restriction of C_5 to any triple.
All 7 compressions must give one of these.
Any $|X| \leq 3$ set is decoded to such a valid triple $Y^{\pm}$ on each $Y \supset X$ with $|Y| = 3$.
E.g., decoding of $X = \{2\}$ cannot give $\{1^-, 2^-, 3^-\}$, $\{0^+, 2^+, 4^+\}$ etc.
It also cannot be $\{1^-, 2^+, 3^-\}$ because some $X \subset \{1, 3\}$ has to give $\{1^-, 3^-\}$.
Thus $\{2\}$ is decoded to $\{0^-, 1^+, 2^?, 3^+, 4^-\}$.
If $\{2\}$ is decoded to $\{0^-, 1^+, 2^+, 3^+, 4^-\}$ and $\{1\}$ is decoded to $\{0^+, 1^+, 2^+, 3^-, 4^-\}$, then they clash at $\{1, 2, 4\}$.
If $\{2\}$ is decoded to $\{0^-, 1^+, 2^-, 3^+, 4^-\}$ and $\{0\}$ is decoded to $\{0^-, 1^+, 2^-, 3^-, 4^+\}$, then they clash at $\{0, 1, 2\}$.

$|C_5| = 10$, the 5 rotations of the following sets:



One of these must occur somewhere. □

Joins

Can we amplify the gap $\mathrm{VC} = 2$ and $\mathrm{UCS} = 3$?

Joins

Can we amplify the gap $\mathrm{VC} = 2$ and $\mathrm{UCS} = 3$?

Definition. Join of two families $\mathcal{F} * \mathcal{G} = \{F \dot\cup G \mid F \in \mathcal{F}, G \in \mathcal{G}\}$.

Joins

Can we amplify the gap $\mathrm{VC} = 2$ and $\mathrm{UCS} = 3$?

Definition. Join of two families $\mathcal{F} * \mathcal{G} = \{F \dot\cup G \mid F \in \mathcal{F}, G \in \mathcal{G}\}$.

In $\pm$ sequence terminology $\mathcal{F} * \mathcal{G} = \{\pm_F {}^\frown \pm_G \mid \pm_F \in \mathcal{F}, \pm_G \in \mathcal{G}\}$.

Joins

Can we amplify the gap $\mathrm{VC} = 2$ and $\mathrm{UCS} = 3$?

Definition. Join of two families $\mathcal{F} * \mathcal{G} = \{F \dot\cup G \mid F \in \mathcal{F}, G \in \mathcal{G}\}$.

In $\pm$ sequence terminology $\mathcal{F} * \mathcal{G} = \{\pm_F {}^\frown \pm_G \mid \pm_F \in \mathcal{F}, \pm_G \in \mathcal{G}\}$.

$\mathrm{VC}(\mathcal{F} * \mathcal{G}) = \mathrm{VC}(\mathcal{F}) + \mathrm{VC}(\mathcal{G})$.

Joins

Can we amplify the gap $\mathrm{VC} = 2$ and $\mathrm{UCS} = 3$?

Definition. Join of two families $\mathcal{F} * \mathcal{G} = \{F \dot\cup G \mid F \in \mathcal{F}, G \in \mathcal{G}\}$.

In $\pm$ sequence terminology $\mathcal{F} * \mathcal{G} = \{\pm_F{}^\frown\pm_G \mid \pm_F \in \mathcal{F}, \pm_G \in \mathcal{G}\}$.

$\mathrm{VC}(\mathcal{F} * \mathcal{G}) = \mathrm{VC}(\mathcal{F}) + \mathrm{VC}(\mathcal{G})$.

But $\mathrm{UCS}(C_5 * C_5) \leq 5 < \mathrm{UCS}(C_5) + \mathrm{UCS}(C_5)$:

Joins

Can we amplify the gap $\mathrm{VC} = 2$ and $\mathrm{UCS} = 3$?

Definition. Join of two families $\mathcal{F} * \mathcal{G} = \{F \dot{\cup} G \mid F \in \mathcal{F}, G \in \mathcal{G}\}$.

In $\pm$ sequence terminology $\mathcal{F} * \mathcal{G} = \{\pm_F {}^\frown \pm_G \mid \pm_F \in \mathcal{F}, \pm_G \in \mathcal{G}\}$.

$\mathrm{VC}(\mathcal{F} * \mathcal{G}) = \mathrm{VC}(\mathcal{F}) + \mathrm{VC}(\mathcal{G})$.

But $\mathrm{UCS}(C_5 * C_5) \leq 5 < \mathrm{UCS}(C_5) + \mathrm{UCS}(C_5)$:

Sample	*Compression*	*Decoding*
no triples	mark all $+$'s	marks $+$, rest -
triple $+$ in $\mathcal{F}$	mark triple and $+$'s in $\mathcal{G}$	triple from position, marks in $\mathcal{G}$ same
triple - in $\mathcal{F}$	mark triple and -'s in $\mathcal{G}$	
triple $+$ in $\mathcal{G}$	mark triple and -'s in $\mathcal{F}$	triple from position, marks in $\mathcal{F}$ opposite
triple - in $\mathcal{G}$	mark triple and $+$'s in $\mathcal{F}$	

Joins

Even stranger is $\mathrm{UCS}(C_5 * [1]) = 3$:

Joins

Even stranger is $\mathrm{UCS}(C_5 * [1]) = 3$:

Sample	*Compression*	*Decoding*
[1] is not -	mark $+$'s of $\mathcal{F}$	marks and [1] $+$, others -
[1] is -, no triple -	mark [1] and -'s	marks and [1] -, others $+$
triple - and [1] is -	mark triple -	marks and [1] -, others $+$

Joins

Even stranger is $\mathrm{UCS}(C_5 * [1]) = 3$:

Sample	*Compression*	*Decoding*
[1] is not -	mark $+$'s of $\mathcal{F}$	marks and [1] $+$, others -
[1] is -, no triple -	mark [1] and -'s	marks and [1] -, others $+$
triple - and [1] is -	mark triple -	marks and [1] -, others $+$

Corollary.
$\mathrm{UCS}(C_5 * C_5) \leq \mathrm{UCS}(C_5 * [1]) + \mathrm{UCS}(C_5 \setminus [1]) = 3 + 2 = 5.$

Joins

Even stranger is $\mathrm{UCS}(C_5 * [1]) = 3$:

Sample	*Compression*	*Decoding*
[1] is not -	mark $+$'s of $\mathcal{F}$	marks and [1] $+$, others -
[1] is -, no triple -	mark [1] and -'s	marks and [1] -, others $+$
triple - and [1] is -	mark triple -	marks and [1] -, others $+$

Corollary.
$\mathrm{UCS}(C_5 * C_5) \leq \mathrm{UCS}(C_5 * [1]) + \mathrm{UCS}(C_5 \setminus [1]) = 3 + 2 = 5.$

Open. Is $\mathrm{UCS}(\mathcal{F} * [n]) = \mathrm{VC}(\mathcal{F} * [n])$ for large enough $n = n(\mathcal{F})$?

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}.$

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}$.

What happens to $\mathrm{UCS}(\mathcal{F}^{*n})$ when $n \to \infty$?

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}$.

What happens to $\mathrm{UCS}(\mathcal{F}^{*n})$ when $n \to \infty$?

$\mathrm{UCS}(C_5^{*5}) \leq 5 \cdot \mathrm{VC}(C_5) + 1 = 11.$

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}$.

What happens to $\mathrm{UCS}(\mathcal{F}^{*n})$ when $n \to \infty$?

$\mathrm{UCS}(C_5^{*5}) \leq 5 \cdot \mathrm{VC}(C_5) + 1 = 11.$

$\mathrm{UCS}(C_5^{*n}) = 2n + 1$ or $\approx 2n + \Theta(\log n)$ or $\approx (2 + \delta)n$?

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}$.

What happens to $\mathrm{UCS}(\mathcal{F}^{*n})$ when $n \to \infty$?

$\mathrm{UCS}(C_5^{*5}) \leq 5 \cdot \mathrm{VC}(C_5) + 1 = 11$.

$\mathrm{UCS}(C_5^{*n}) = 2n + 1$ or $\approx 2n + \Theta(\log n)$ or $\approx (2 + \delta)n$?

How to get lower bounds?

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}.$

What happens to $\mathrm{UCS}(\mathcal{F}^{*n})$ when $n \to \infty$?

$\mathrm{UCS}(C_5^{*5}) \leq 5 \cdot \mathrm{VC}(C_5) + 1 = 11.$

$\mathrm{UCS}(C_5^{*n}) = 2n + 1$ or $\approx 2n + \Theta(\log n)$ or $\approx (2 + \delta)n$?

How to get lower bounds?

Even possible that $\mathrm{UCS}(C_5 * C_5) = 4$!

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}$.

What happens to $\mathrm{UCS}(\mathcal{F}^{*n})$ when $n \to \infty$?

$\mathrm{UCS}(C_5^{*5}) \leq 5 \cdot \mathrm{VC}(C_5) + 1 = 11$.

$\mathrm{UCS}(C_5^{*n}) = 2n + 1$ or $\approx 2n + \Theta(\log n)$ or $\approx (2 + \delta)n$?

How to get lower bounds?

Even possible that $\mathrm{UCS}(C_5 * C_5) = 4$!

Is $\mathrm{UCS}(\mathcal{F} * \mathcal{G}) \geq \mathrm{UCS}(\mathcal{F}) + \mathrm{UCS}(\mathcal{G}) - 1$?

Limits

Definition. $\mathcal{F}^{*n} = \underbrace{\mathcal{F} * \ldots * \mathcal{F}}_{n \text{ times}}$.

What happens to $\mathrm{UCS}(\mathcal{F}^{*n})$ when $n \to \infty$?

$\mathrm{UCS}(C_5^{*5}) \leq 5 \cdot \mathrm{VC}(C_5) + 1 = 11$.

$\mathrm{UCS}(C_5^{*n}) = 2n + 1$ or $\approx 2n + \Theta(\log n)$ or $\approx (2 + \delta)n$?

How to get lower bounds?

Even possible that $\mathrm{UCS}(C_5 * C_5) = 4$!

Is $\mathrm{UCS}(\mathcal{F} * \mathcal{G}) \geq \mathrm{UCS}(\mathcal{F}) + \mathrm{UCS}(\mathcal{G}) - 1$?

$\mathrm{UCS}(W_6 * W_6) \leq 2 \cdot \mathrm{VC}(W_6) + 1 = 5$.

Occasional One-sided Getback

Completely new game:

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots,$

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not.

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err,

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err, but even in this case only occasionally.

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err, but even in this case only occasionally.

$$\mathrm{O1G}(\mathcal{F}) = \min_{\substack{(x_i) \text{ ordering and} \\ \text{guessing strategy}}} \max_{\substack{F \in \mathcal{F} \text{ and} \\ \text{getback strategy}}} \#\textit{error messages.}$$

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err, but even in this case only occasionally.

$$\mathrm{O1G}(\mathcal{F}) = \min_{\substack{(x_i)\ \textit{ordering and}\\ \textit{guessing strategy}}} \max_{\substack{F\in\mathcal{F}\ \textit{and}\\ \textit{getback strategy}}} \#\textit{error messages}.$$

Claim. $\mathrm{UCS} \leq \mathrm{O1G}$.

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err, but even in this case only occasionally.

$$\mathrm{O1G}(\mathcal{F}) = \min_{\substack{(x_i)\ \textit{ordering and} \\ \textit{guessing strategy}}} \max_{\substack{F\in\mathcal{F}\ \textit{and} \\ \textit{getback strategy}}} \#\textit{error messages.}$$

Claim. $\mathrm{UCS} \leq \mathrm{O1G}$.

Proof. One-Pass Compression Scheme (Floyd-Warmuth) by marking where algorithm would err on sample.

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err, but even in this case only occasionally.

$$\mathrm{O1G}(\mathcal{F}) = \min_{\substack{(x_i) \text{ ordering and} \\ \text{guessing strategy}}} \max_{\substack{F \in \mathcal{F} \text{ and} \\ \text{getback strategy}}} \#\textit{error messages.}$$

Claim. $\mathrm{UCS} \leq \mathrm{O1G}$.

Proof. One-Pass Compression Scheme (Floyd-Warmuth) by marking where algorithm would err on sample.
Decoding straight-forward. □

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err, but even in this case only occasionally.

$$\mathrm{O1G}(\mathcal{F}) = \min_{\substack{(x_i) \text{ ordering and} \\ \text{guessing strategy}}} \max_{\substack{F \in \mathcal{F} \text{ and} \\ \text{getback strategy}}} \#\textit{error messages.}$$

Claim. $\mathrm{UCS} \leq \mathrm{O1G}$.

Proof. One-Pass Compression Scheme (Floyd-Warmuth) by marking where algorithm would err on sample.
Decoding straight-forward. □

$\mathrm{O1G}(\mathcal{F} * \mathcal{G}) = \mathrm{O1G}(\mathcal{F}) + \mathrm{O1G}(\mathcal{G})$.

Occasional One-sided Getback

Completely new game: Elements come one by one, $x_1, x_2, \ldots$, each time we have to guess whether $x_i \in F$ or not. Only info we get back is if we err, but even in this case only occasionally.

$$\mathrm{O1G}(\mathcal{F}) = \min_{\substack{(x_i) \text{ ordering and} \\ \text{guessing strategy}}} \max_{\substack{F \in \mathcal{F} \text{ and} \\ \text{getback strategy}}} \#\textit{error messages.}$$

Claim. $\mathrm{UCS} \leq \mathrm{O1G}$.

Proof. One-Pass Compression Scheme (Floyd-Warmuth) by marking where algorithm would err on sample.
Decoding straight-forward. □

$\mathrm{O1G}(\mathcal{F} * \mathcal{G}) = \mathrm{O1G}(\mathcal{F}) + \mathrm{O1G}(\mathcal{G})$.

Open. Is $\mathrm{O1G}(\mathcal{F}) = O(\mathrm{VC}(\mathcal{F}))$ for finite $\mathcal{F}$?

Thank you for your attention!