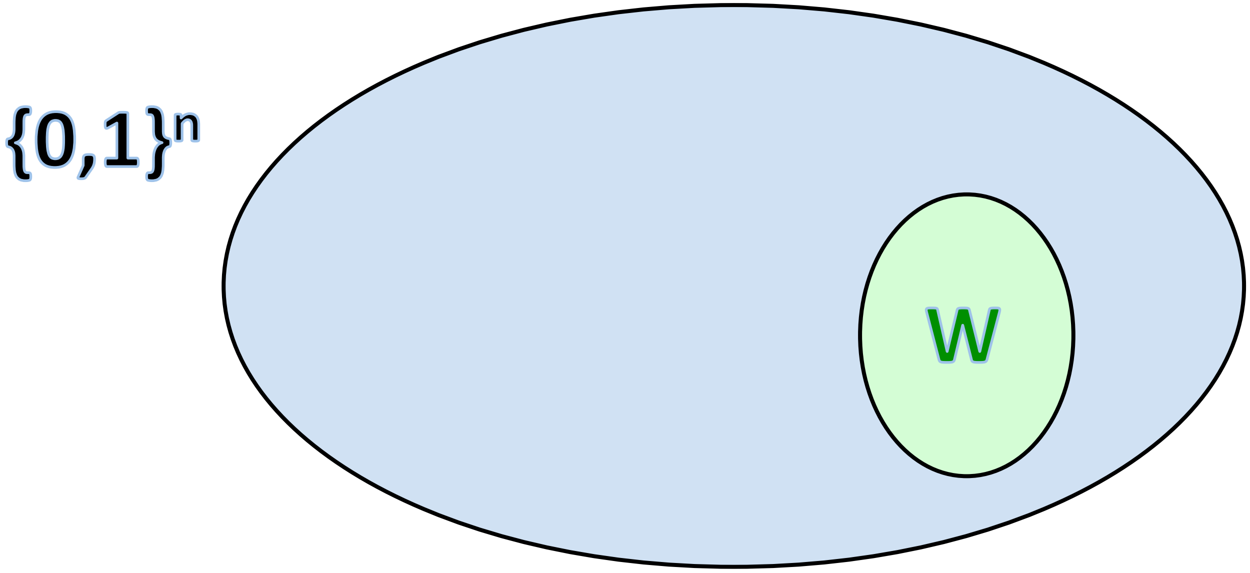
The Query Complexity of Witness Finding

Akinori Kawachi Ben Rossman Osamu Watanabe
Tokyo Tech NII Tokyo Tech

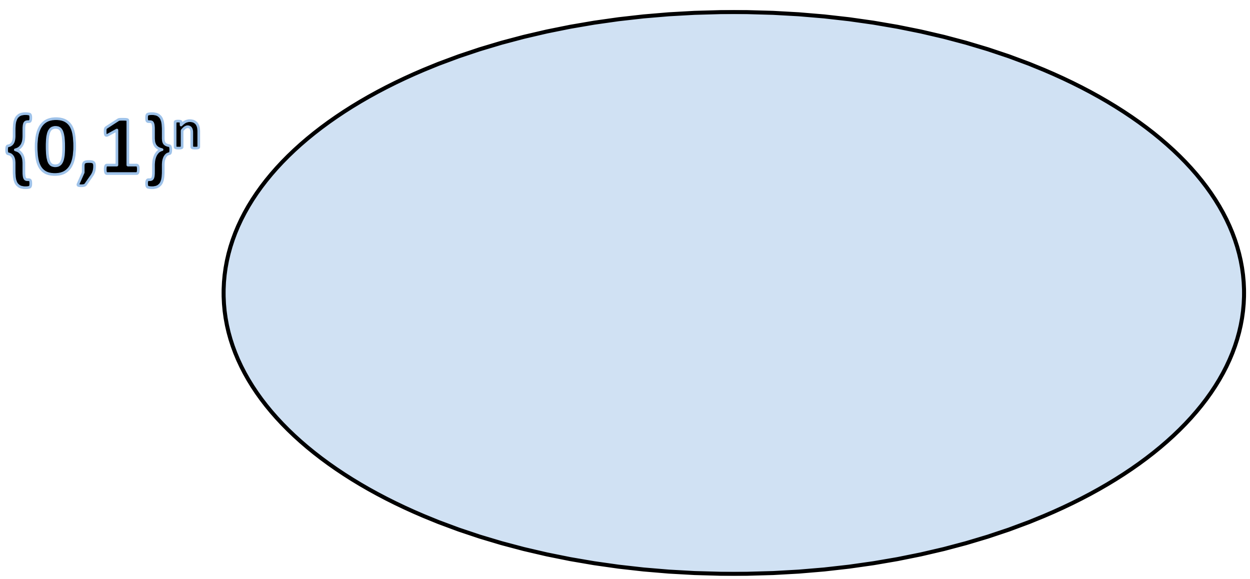
Witness Finding Problem

nonempty **witness set** W ⊆ $\{0,1\}^n$



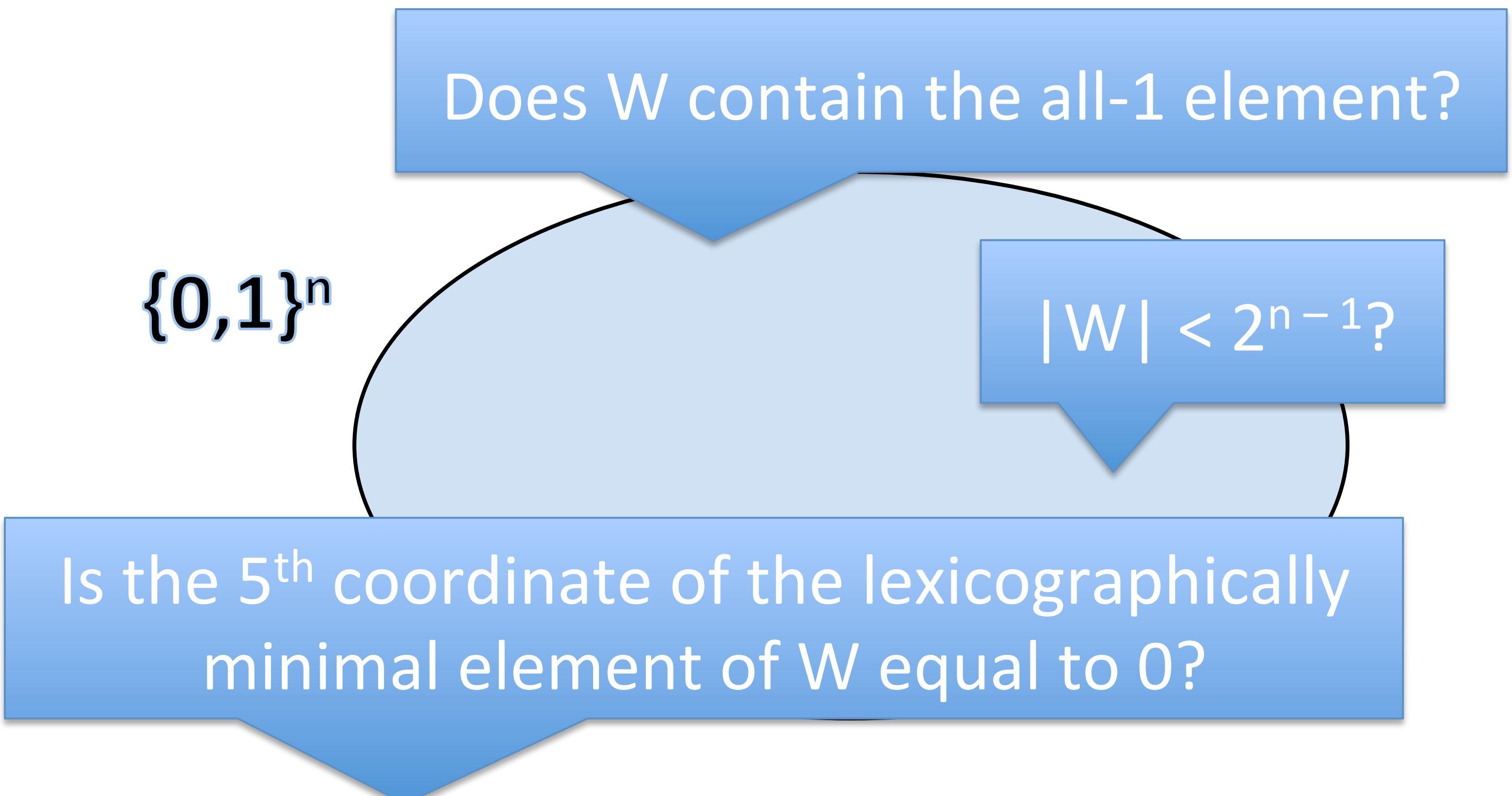
Witness Finding Problem

- Hidden nonempty **witness set** $W \subseteq \{0,1\}^n$



Witness Finding Problem

- <u>Hidden</u> nonempty **witness set** $W \subseteq \{0,1\}^n$
- We ask **queries** (yes/no question about W)

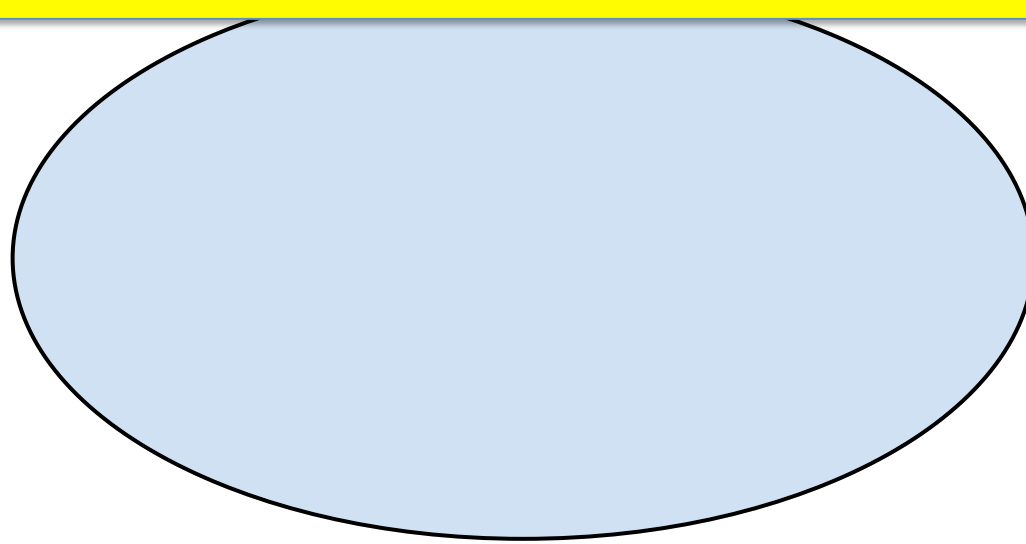


Witness Finding Problem

- Hidden nonempty **witness set** W $\subseteq \{0,1\}^n$
- We ask **queries** (yes/no question about W)

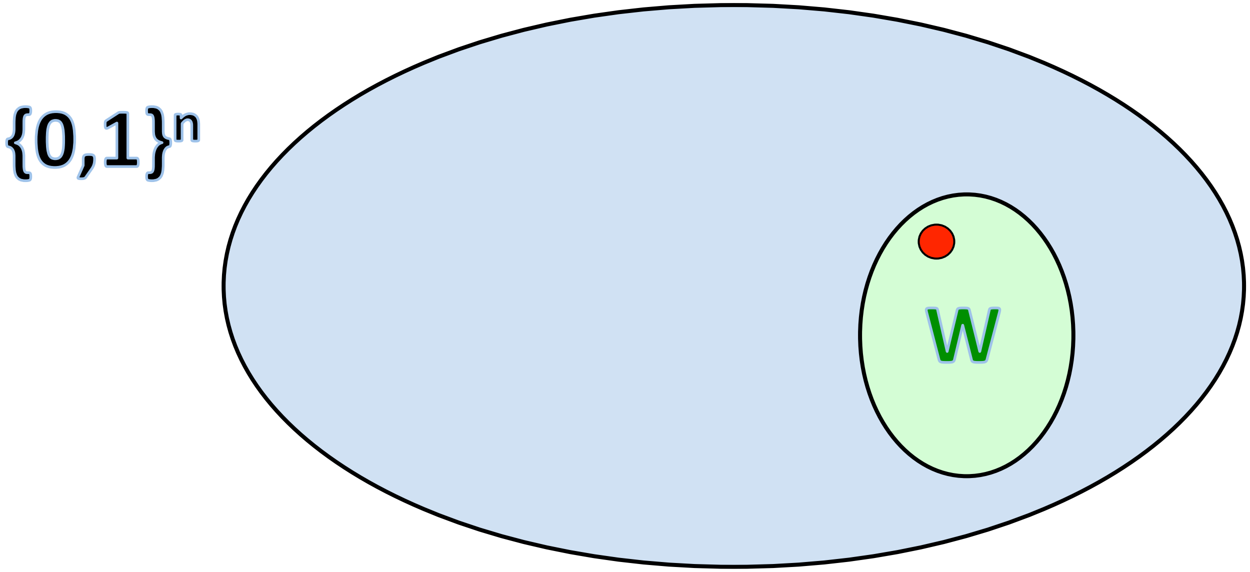
Queries are ***randomized*** and ***non-adaptive.***

$\{0,1\}^n$



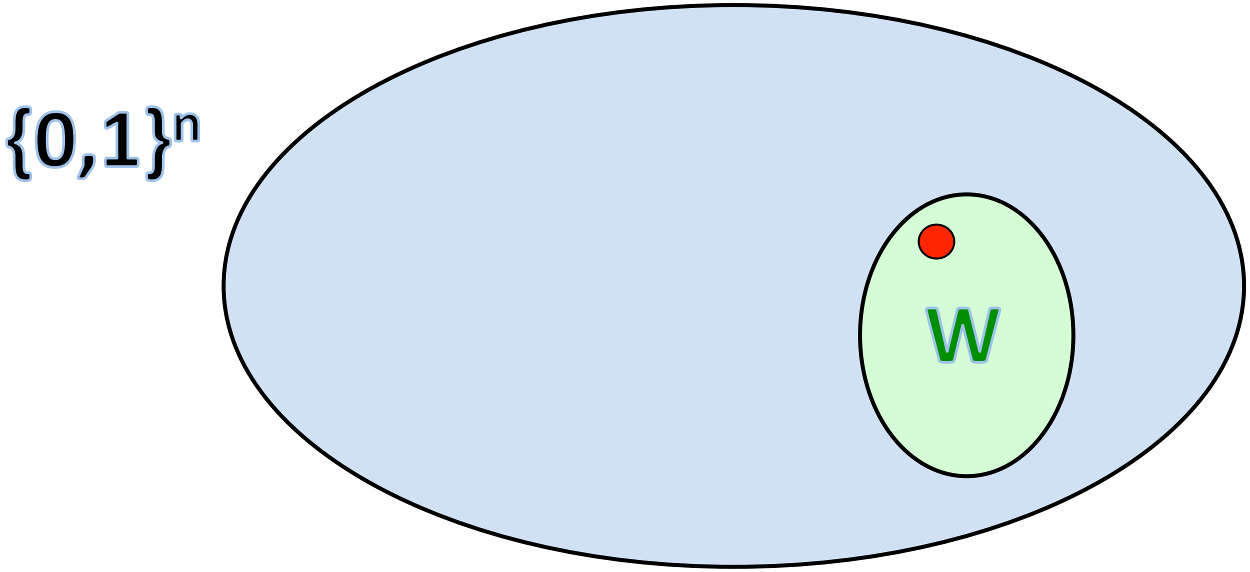
Witness Finding Problem

- After receiving yes/no answers to our queries, we output an element $\mathbf{x \in \{0,1\}^n}$
- We succeed iff $\mathbf{x \in W}$



Witness Finding Problem

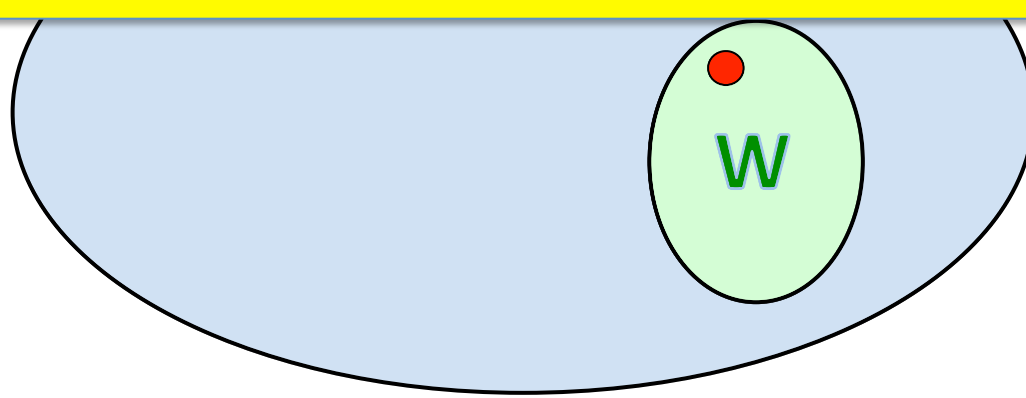
- Goal: Succeed with probability $> 1/2$ for *every* nonempty $W \subseteq \{0,1\}^n$



Witness Finding Problem

- <u>Goal</u>: Succeed with probability $> 1/2$ for *every* nonempty $W \subseteq \{0,1\}^n$

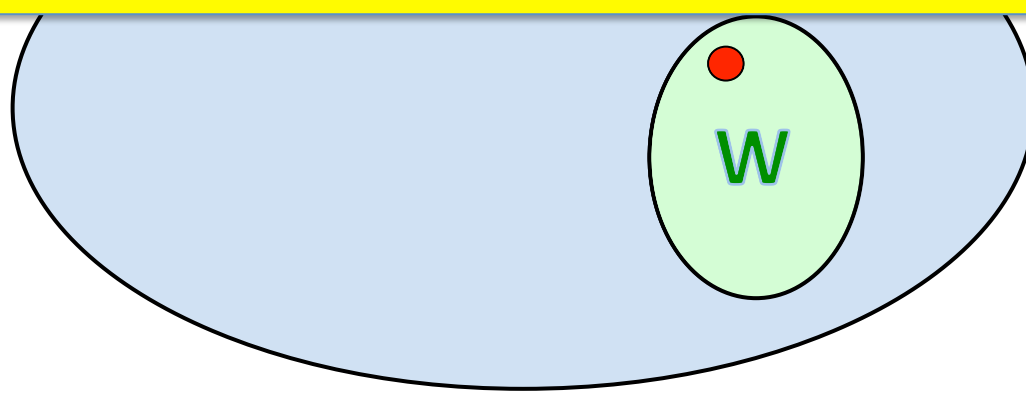
We are interested in the ***query complexity*** of this problem: the fewest number of (non-adaptive, randomized) queries



Witness Finding Problem

- Goal: Succeed with probability $> 1/2$ for *every* nonempty $W \subseteq \{0,1\}^n$

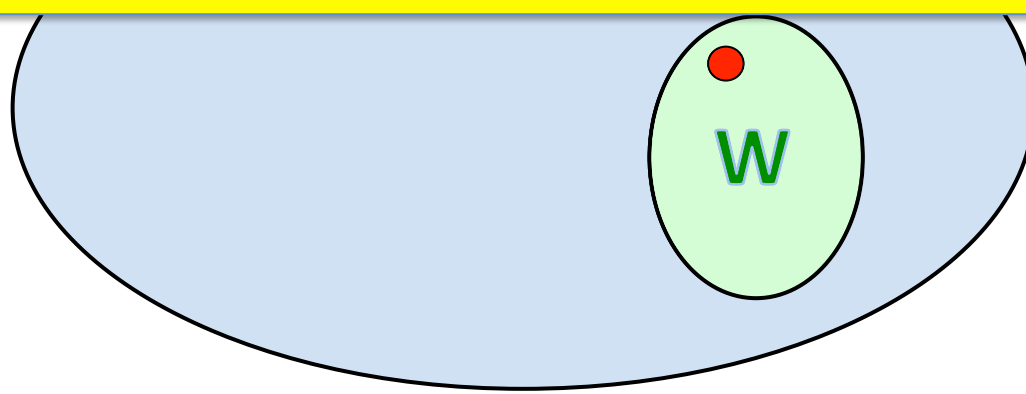
We are interested in the ***query complexity*** of this problem: the fewest number of (non-adaptive, randomized) queries **from a specific class of permitted queries**



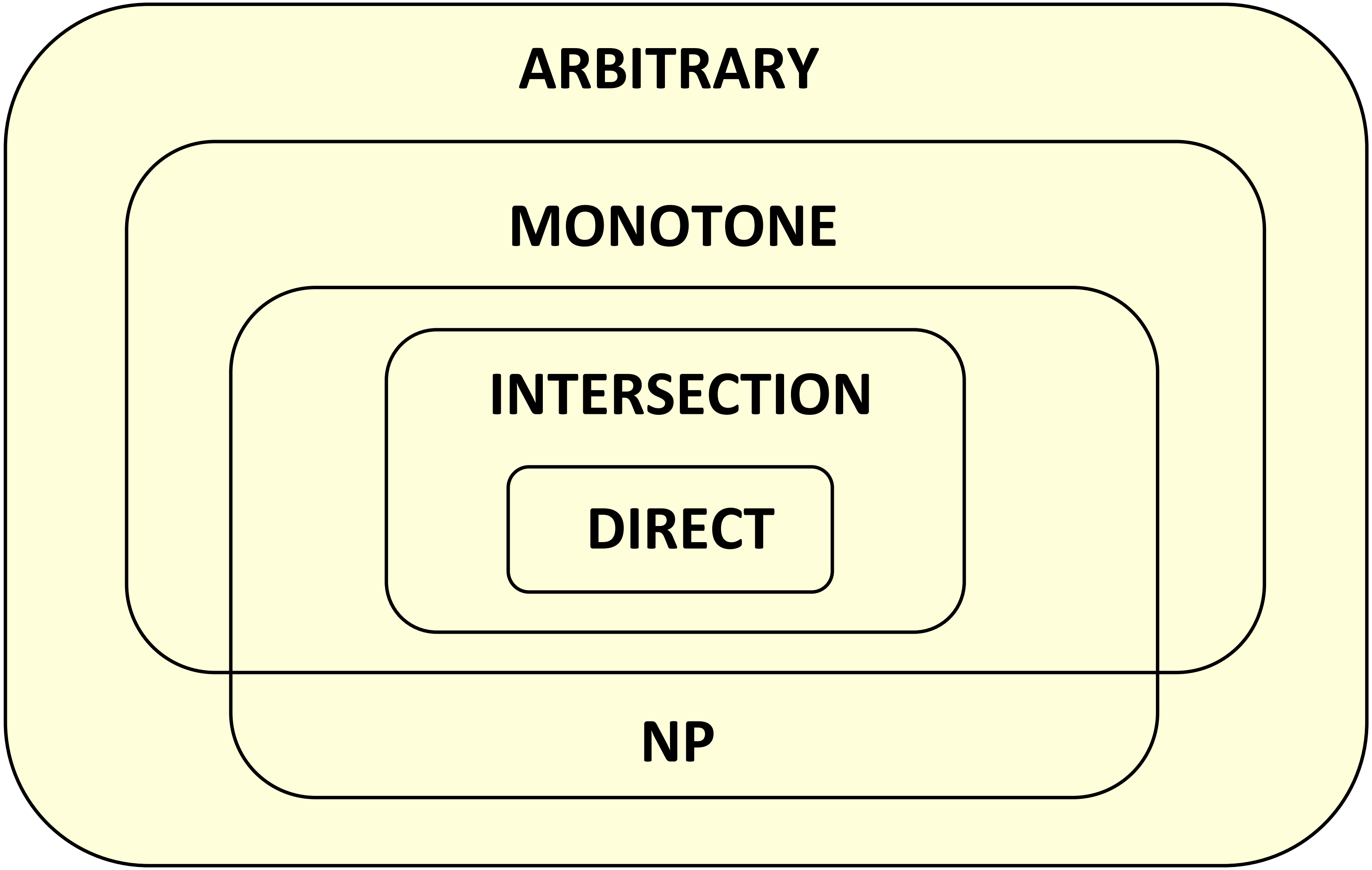
Witness Finding Problem

- Goal: Succeed with probability $> 1/2$ for *every* nonempty $W \subseteq \{0,1\}^n$

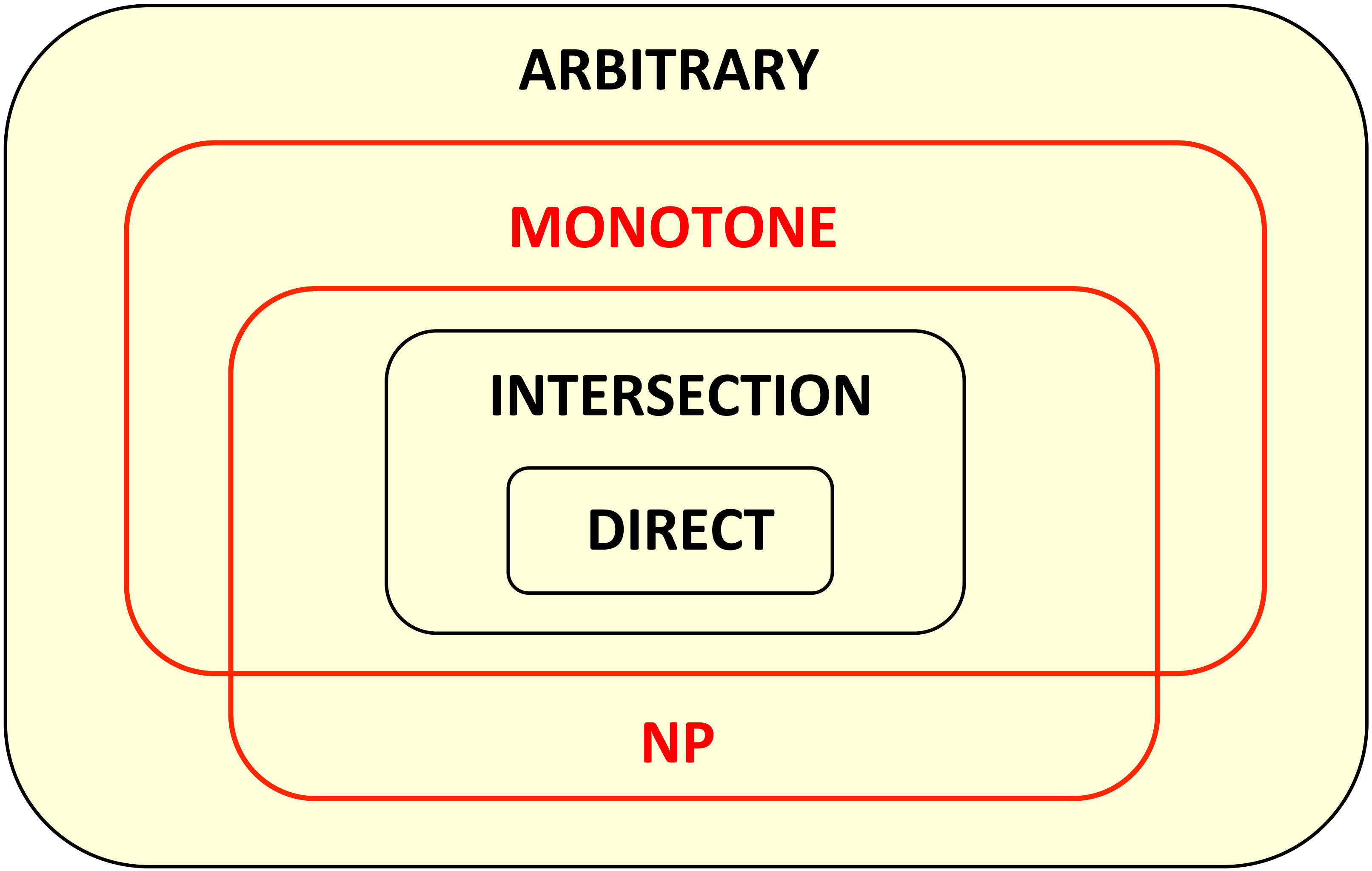
Our results are tight, information-theoretic **lower bounds** on the query complexity of witness finding for a few natural classes of queries.



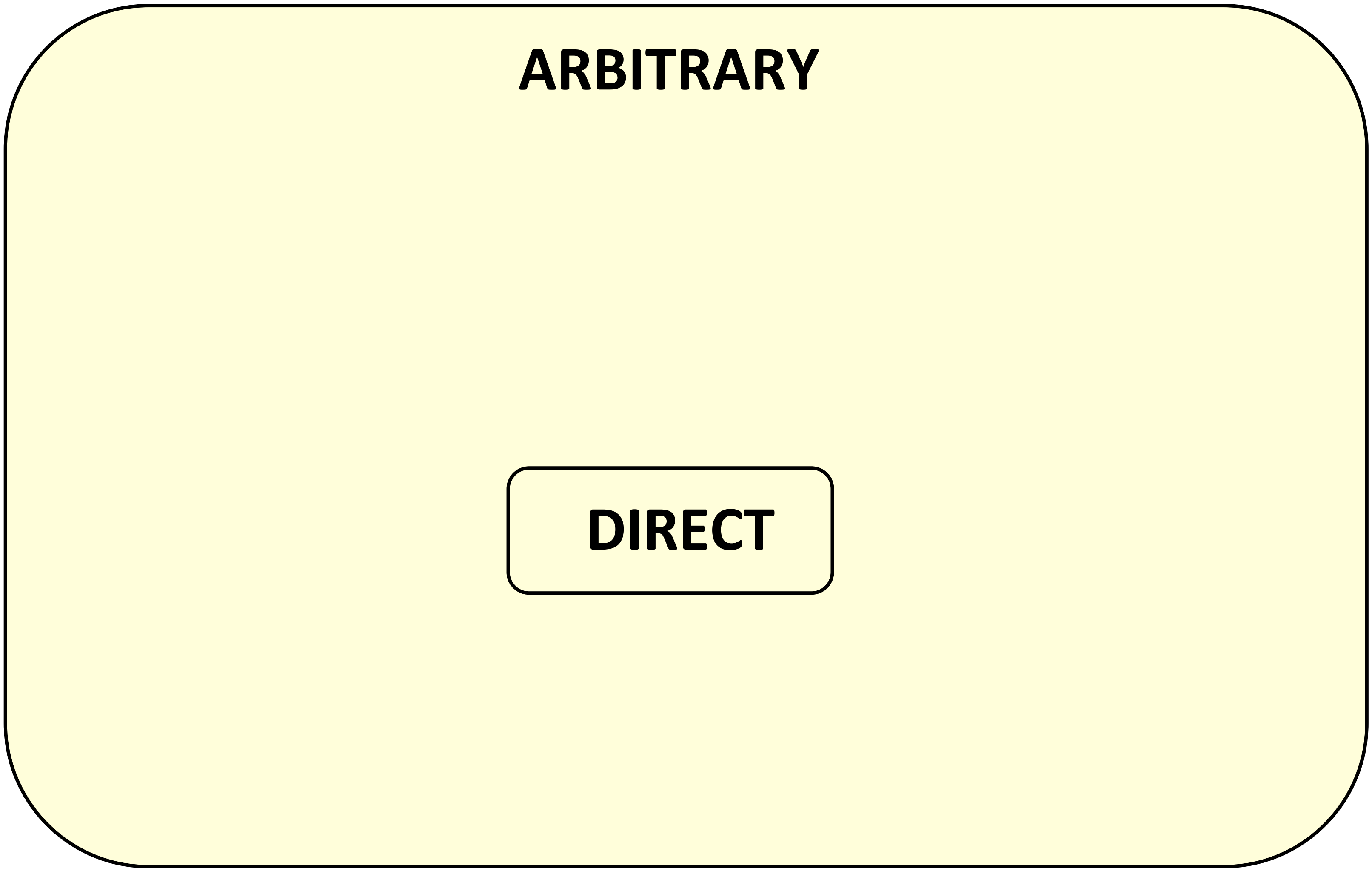
Classes of Queries



Classes of Queries



Two Trivial Classes



Two Trivial Classes

DIRECT QUERY

"Is $x \in W$?" where $x \in \{0,1\}^n$

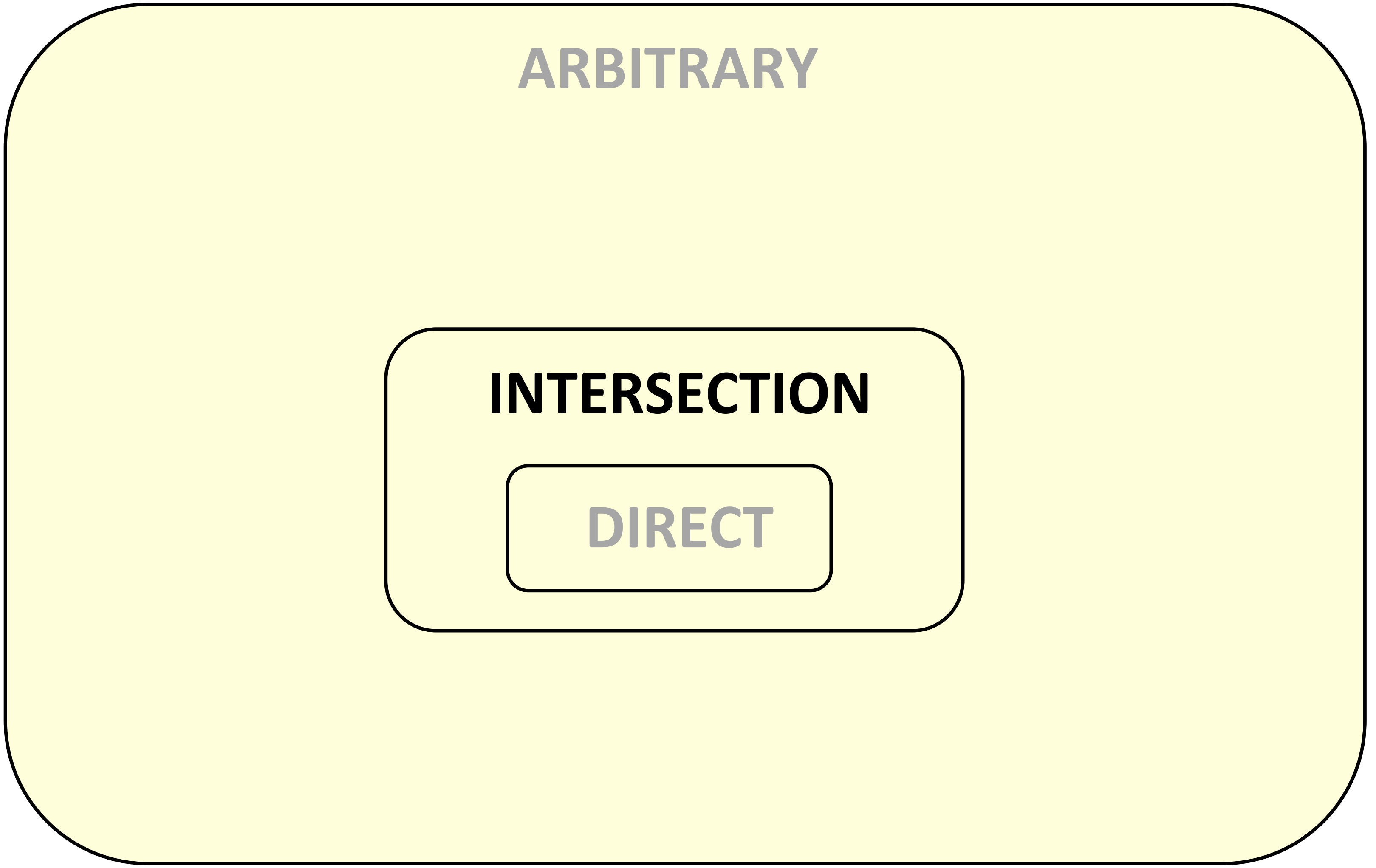
2^n direct queries are necessary and sufficient to find a witness in every W with probability $> ½$

ARBITRARY QUERY

"Is $W \in F$?" where $F \subseteq Pow(\{0,1\}^n)$

n arbitrary queries are necessary and sufficient

Intersection Queries



Intersection Queries

INTERSECTION QUERY

"Is $S \cap W$ nonempty?" where $S \subseteq \{0,1\}^n$

Theorem (Ben-David, Chor, Goldreich, Luby)

Witness finding is solvable with $O(n^2)$ intersection queries.

We show

Witness finding requires $\Omega(n^2)$ intersection queries.

Theorem (Ben-David, Chor, Goldreich, Luby)

Witness finding is solvable with $O(n^2)$ intersection queries.

- Uses *Valiant-Vazirani Isolation Lemma*.

Theorem (Ben-David, Chor, Goldreich, Luby)

Witness finding is solvable with $O(n^2)$ intersection queries.

- If we know that $2^k \le |W| \le 2^{k+1}$, then $O(n)$ intersection queries suffice ($\Rightarrow O(n^2)$ upper bound)
- For random $S \subseteq \{0,1\}^n$ of density 2^{-k}, $|S \cap W| = 1$ with constant probability (> 1/100).
- With $2 \log|S|$ $(= O(n))$ simultaneous intersection queries, we can detect whether $|S \cap W| = 1$ and identify the unique element: if $S = \{x_1,\ldots,x_{|S|}\}$, we ask

  "does W intersect $\{ x_i \mid t^{th}$ bit of i equals b }?"

  for all $i \in \{1,\ldots,\log|S|\}$ and $b \in \{0,1\}$

Theorem (Ben-David, Chor, Goldreich, Luby)

Witness finding is solvable with $O(n^2)$ intersection queries.

- Gives an $\mathbf{BPP}_{\mathbf{||}}^{\mathbf{NP}}$ algorithm (***search-to-decision reduction***) that solves Search(Circuit-SAT) by making $O(n^2)$ non-adaptive calls to an oracle for Decision(Circuit-SAT).

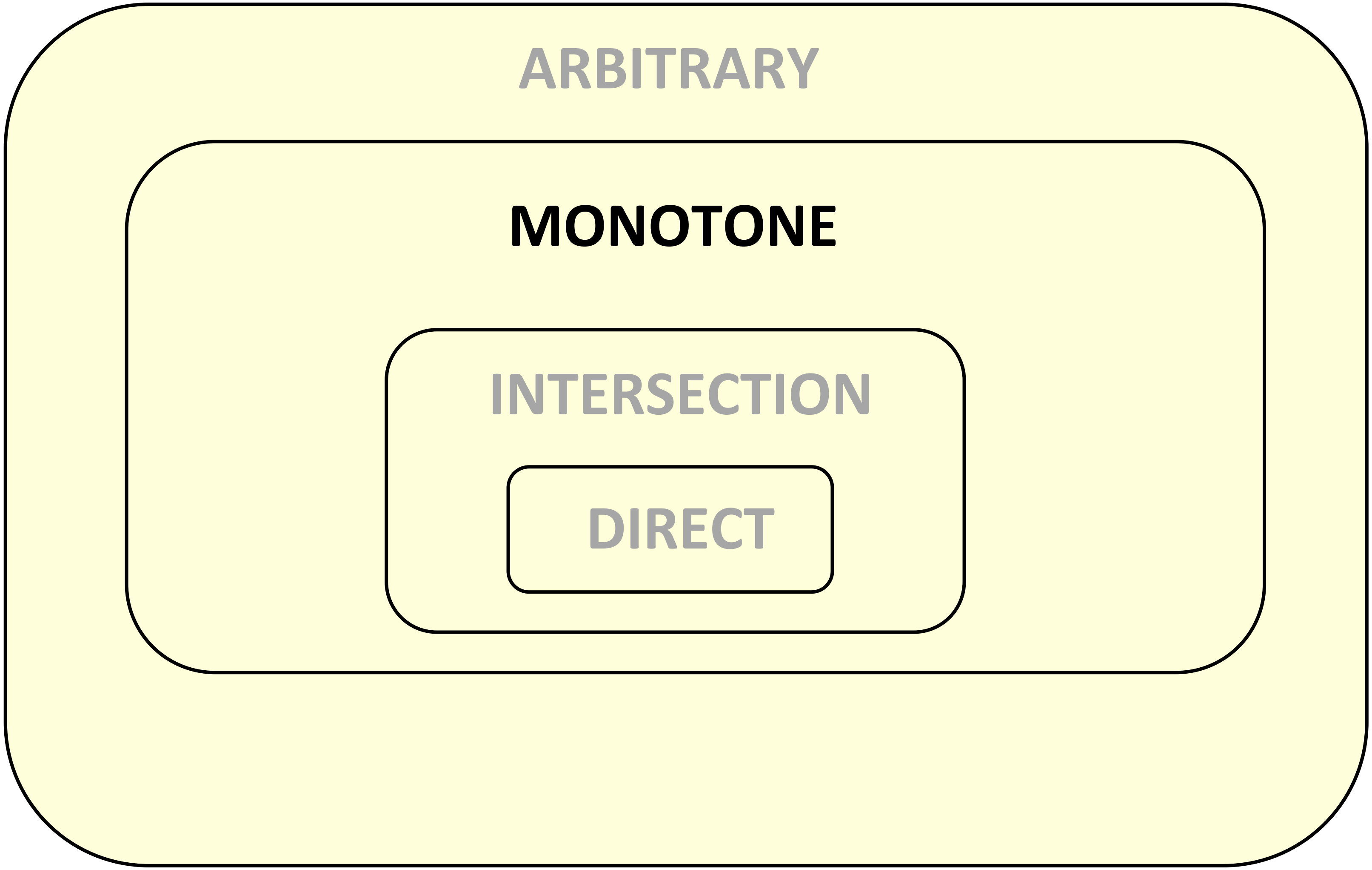
Theorem (Ben-David, Chor, Goldreich, Luby)

Witness finding is solvable with $O(n^2)$ intersection queries.

- Gives an $\mathbf{BPP}_{||}^{\mathbf{NP}}$ algorithm (***search-to-decision reduction***) that solves Search(Circuit-SAT) by making $O(n^2)$ non-adaptive calls to an oracle for Decision(Circuit-SAT).

Obs: This search-to-decision reduction is **black-box:** it never "looks" at the input circuit C; it merely requires an oracle to the witness set $\{x \mid C(x) = 1\}$.

Monotone Queries



Monotone Queries

An **monotone query** is a query of the form

"$f(W) = 1$?"

where $f : Pow(\{0,1\}^n) \longrightarrow \{0,1\}$ is a monotone function

- Every intersection query is monotone.

Monotone Queries

An **monotone query** is a query of the form

"$f(W) = 1$?"

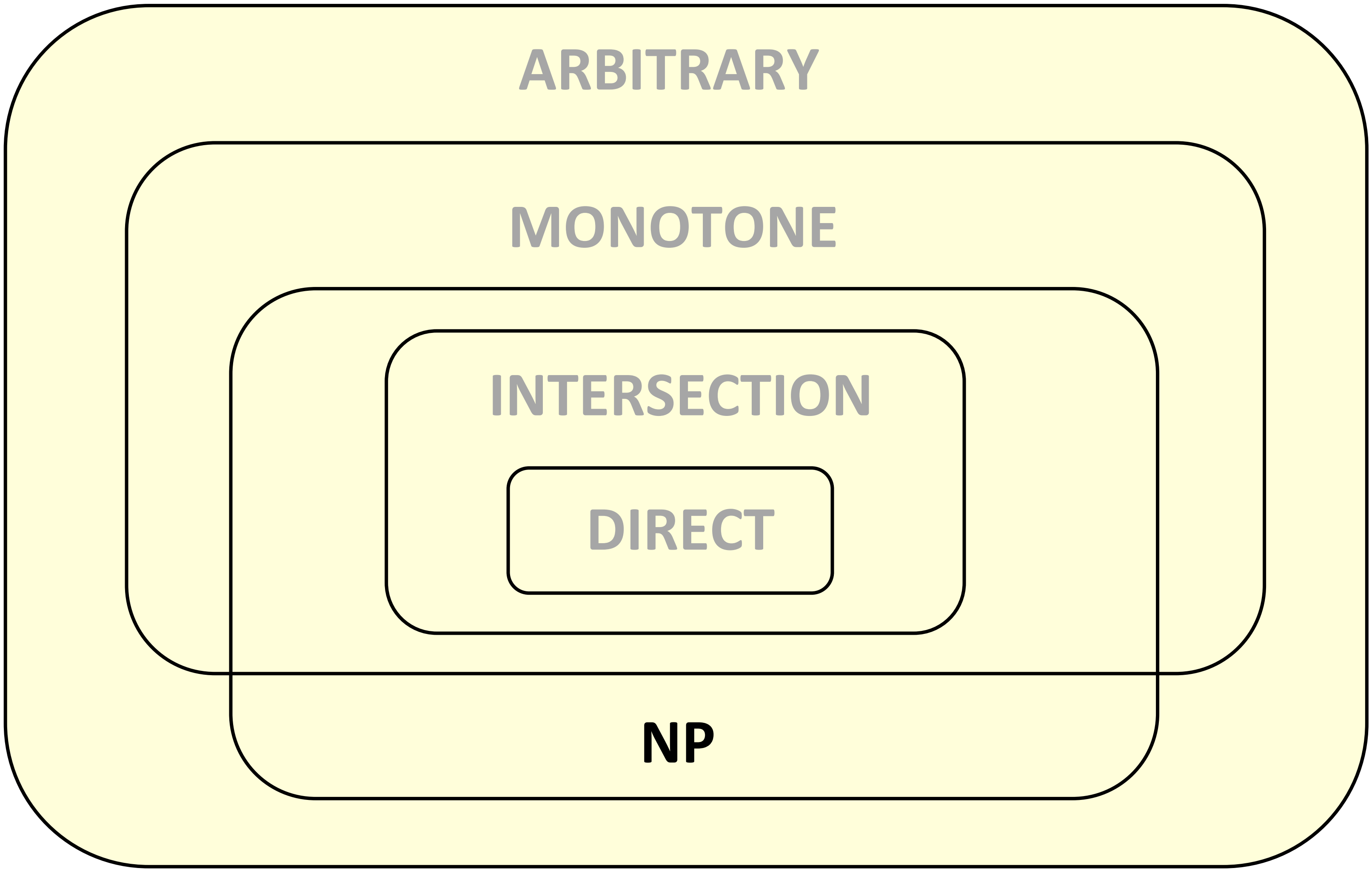
where $f : \mathrm{Pow}(\{0,1\}^n) \longrightarrow \{0,1\}$ is a monotone function

- Every intersection query is monotone.

Theorem

Witness finding requires $\Omega(n^2)$ monotone queries.

NP Queries



NP Queries

An **NP query** is a query of the form

"A(W) = 1?"

where A is a fixed ***non-deterministic algorithm*** which makes ***poly(n) direct queries*** and outputs a single bit

- A can guess a witness in W. However, A cannot guess the lexicographically minimal element of W.

NP Queries

An **NP query** is a query of the form

"A(W) = 1?"

where A is a fixed ***non-deterministic algorithm*** which makes ***poly(n) direct queries*** and outputs a single bit

- NP queries not necessarily monotone (& vice-versa)
- However, every intersection query is an NP query: given $S \subseteq \{0,1\}^n$, non-deterministically guess $x \in S$ and verify that $x \in W$ using a single direct query.

NP Queries

Main Theorem

Witness finding requires $\Omega(n^2)$ NP queries.

- This shows that the procedure of Ben-David et al. has optimal query complexity among ***black-box*** ***$BPP_{||}^{NP}$ search-to-decision reductions***.

PROOF SKETCHES

Theorem

Witness finding requires $\Omega(n^2)$ intersection queries.

Theorem

Witness finding requires $\Omega(n^2)$ intersection queries.

- Want a lower bound on ***randomized algorithms*** which output an element of W with probability > ½ for ***every fixed*** nonempty witness set $W \subseteq \{0,1\}^n$.
- Invoking Yao's Minimax Principle, we flip the situation: we fix a ***distribution*** on witness sets and show that every ***deterministic algorithm*** which succeeds on this distribution with probability > ½ requires $\Omega(n^2)$ intersection queries.

- We define the ***distribution on W*** as follows:

1. pick $K \in \{1,\ldots,n\}$ uniform at random,
2. pick W uniformly among subsets of $\{0,1\}^n$ of size 2^K

- We define the ***distribution on W*** as follows:
  1. pick $K \in \{1,\ldots,n\}$ uniform at random,
  2. pick W uniformly among subsets of $\{0,1\}^n$ of size 2^K

- Using this same distribution, Dell, Kabanets, van Melkebeek, Watanabe [CCC'12] proved an $O(1/n)$ upper bound on the ***success probability*** of black-box ***witness-isolation*** procedures.

- A ***deterministic witness finding algorithm with m intersection queries*** is specified by

$$S_1, \ldots, S_m \subseteq \{0,1\}^n$$

$$f : \{0,1\}^m \longrightarrow \{0,1\}^n$$

  That is, the algorithm:

  1. asks intersection queries "Is $S_i \cap W$ nonempty?"
  2. receives answers $X_1,\ldots,X_m \in \{0,1\}$
  3. outputs $f(X_1,\ldots,X_m) \in \{0,1\}^n$

- We view $X_1,\ldots,X_m$ as ***0-1 valued random variables*** (completely determined by W, once the algorithm is fixed)

- A ***deterministic witness finding algorithm with m intersection queries*** is specified by

$$S_1, \ldots, S_m \subseteq \{0,1\}^n$$

$$f : \{0,1\}^m \longrightarrow \{0,1\}^n$$

That is, the algorithm:

1. asks intersection queries "Is $S_i \cap W$ nonempty?"
2. receives answers $X_1,\ldots,X_m \in \{0,1\}$
3. outputs $f(X_1,\ldots,X_m) \in \{0,1\}^n$

Theorem (restated)

If $\Pr[\, f(X_1,\ldots,X_m) \in W \,] > ½$, then $m = \Omega(n^2)$

Theorem If $\Pr[\, f(X_1,\ldots,X_m) \in W \,] > ½$, then $m = \Omega(n^2)$

Lemma 1 $H(f(X_1,\ldots,X_m)) = \Omega(n)$

Lemma 2 $H(X_i \mid K) = O(1/n)$ for every i

Theorem If $\Pr[\, f(X_1,\ldots,X_m) \in W \,] > ½$, then $m = \Omega(n^2)$

Lemma 1 $H(f(X_1,\ldots,X_m)) = \Omega(n)$

Lemma 2 $H(X_i \mid K) = O(1/n)$ for every i

Proof of Lemmas 1&2 => Theorem:

$$\begin{aligned}\Omega(n) &= H(f(X_1,\ldots,X_m)) \\ &\le H(X_1,\ldots,X_m) \\ &\le H(X_1,\ldots,X_m,K) \\ &= H(K) + H(X_1,\ldots,X_m \mid K) \\ &\le \log(n) + H(X_1 \mid K) + \ldots + H(X_m \mid K) \\ &= \log(n) + O(m/n).\end{aligned}$$

hence, $m = \Omega(n^2)$

Lemma 1 $H(f(X_1,\ldots,X_m)) = \Omega(n)$

- More generally, we show that

  W has ***ε-witness-entropy*** $\Omega(n)$ for every const. $\varepsilon > 0$

  where the ***ε-witness-entropy*** of a random nonempty set U is defined as the minimum H(y) over random variables y such that $\Pr[y \in U] \geq \varepsilon$

- Other examples: The ***uniform random nonempty subset*** of $\{0,1\}^n$ has witness-entropy O(1).

  The ***random affine subspace of $\{0,1\}^n$ of dimension K*** (uniform in {1,…,n}) has ε-witness-entropy $\Omega(n)$ for every $\varepsilon > 0$.

Lemma 2 $H(X_i | K) = O(1/n)$ for every i

- Recall that $X_i \in \{0,1\}$ is the indicator for the event "S_i intersects W" where $S_i \subseteq \{0,1\}^n$

Lemma 2 $H(X_i | K) = O(1/n)$ for every i

- $H(X_i | K) = (1/n) \sum_{k=1}^{n} H(\text{"}S_i \text{ intersects } W\text{"} \mid W \text{ has size } 2^k)$
- Let $t = n - \log |S_i|$ (so $|S_i| = 2^{n-t}$)

Lemma ("k = t is a threshold for X_i")

$k \le t \Rightarrow \Pr[S_i \text{ intersects } W \mid W \text{ has size } 2^k] \le (½)^{\Omega(t-k)}$

$k \ge t \Rightarrow \Pr[S_i \text{ intersects } W \mid W \text{ has size } 2^k] \ge 1 - (½)^{\Omega(k-t)}$

- $H(\text{"}S_i \text{ intersects } W\text{"} \mid W \text{ has size } 2^k) \le (½)^{\Omega(|t-k|)}$
- $H(X_i | K) = (1/n) \sum_{k=1}^{n} (½)^{\Omega(|t-k|)} = O(1/n)$

PROOF SKETCHES

We showed:

Theorem

Witness finding requires $\Omega(n^2)$ intersection queries.

By essentially the same proof, we get:

Theorem

Witness finding requires $\Omega(n^2)$ monotone queries.

Lemma 2' For every monotone $f : \mathrm{Pow}(\{0,1\}^n) \longrightarrow \{0,1\}$, $H(f(W) \mid K) = O(1/n)$.

- For $1 \le k \le n$, let $p_k = E[\, f(W) \mid W \text{ has size } 2^k \,]$
- Assuming f is non-trivial, $0 < p_1 < p_2 < \ldots < p_n = 1$
- Let t be the "threshold" such that $p_t < 1/2 \le p_{t+1}$
- By the ***Bollobas-Thomason Theorem***:

Lemma

$k \le t \;\Rightarrow\; p_k \le (½)^{\Omega(t-k)}$

$k \ge t \;\Rightarrow\; p_k \ge 1 - (½)^{\Omega(k-t)}$

PROOF SKETCHES

Main Theorem

Witness finding requires $\Omega(n^2)$ NP queries.

- Proof by reduction to setting of *monotone* queries: we show that every NP query is ***well-approximated*** by a monotone query.

Lemma

For every NP query Q, there is a monotone query Q^+ such that $\Pr[\, Q(W) \neq Q^+(W) \,] \leq 1/n^{\omega(1)}$

- Q non-deterministically makes poly(n) direct queries and returns a single bit.
- Wlog, Q guesses answers to its queries beforehand and simply verifies.
- We get Q^+ by only verifying answers that are guessed to be positive.

AFFINE SUBSPACES

Too Many Witness Sets?

- For any given NP search problem, there are only $2^{poly(n)}$ possible witness sets.
- In the proof of our lower bounds, the distribution on W has support $2^{exp(n)}$.
- Can a black-box search-to-decision reduction for a specific NP problem (3SAT, say) achieve better than $O(n^2)$ query complexity by exploiting the fact that W is the witness set of some (unseen) 3SAT instance?

Affine Witness Sets

- One natural approach: instead of a ***random subset of {0,1}^{n} of size 2^K*** (where K uniform in {1,…,n}), consider a ***random affine subspace of dimension K***.
- This distribution is the support of an actual NP search problem.

Affine Witness Sets

Theorem

Affine witness finding requires $\Omega(n^2)$ ***intersection*** *queries.*

OPEN

Does affine witness finding require $\Omega(n^2)$ ***monotone*** *queries?*

Let $f : \mathrm{Pow}(\{0,1\}^n) \longrightarrow \{0,1\}$ be a monotone function

- For $1 \le k \le n$, let $p_k = E[\ f(A)\ |\ A \text{ affine space of dim } k\ ]$
- Let t be the "threshold" such that $p_t < 1/2 \le p_{t+1}$

CONJECTURE

$$k \le t \;\Rightarrow\; p_k \le (½)^{\Omega(t-k)}$$

$$k \ge t \;\Rightarrow\; p_k \ge 1 - (½)^{\Omega(k-t)}$$

- We have a proof in the case where f is an ***intersection query*** (i.e. there exists $S \subseteq \{0,1\}^n$ such that $f(A) = 1$ iff A intersects S)

Let $f : \mathrm{Pow}(\{0,1\}^n) \longrightarrow \{0,1\}$ be a monotone function

- For $1 \le k \le n$, let $p_k = \mathrm{E}[\ f(A)\ |\ A \text{ affine space of dim } k\ ]$
- Let t be the "threshold" such that $p_t < 1/2 \le p_{t+1}$

CONJECTURE

$$k \le t \;\Rightarrow\; p_k \le (½)^{\Omega(t-k)}$$

$$k \ge t \;\Rightarrow\; p_k \ge 1 - (½)^{\Omega(k-t)}$$

- There is a "q-analogue" of the ***Bollobas-Thomason Theorem***. However, it merely implies:

$$k \le t \;\Rightarrow\; p_k \le (½)^{\Omega(t/k)}$$

$$k \ge t \;\Rightarrow\; p_k \ge 1 - (½)^{\Omega((n-t)/(n-k))}$$

- Let B(n) = lattice of **subsets** of {1,…,n},

  L(n) = lattice of **linear subspaces** of $\{0,1\}^n$

- On the one hand, L(n) is the "q-analogue" of B(n).

  On the other hand, L(n) is a sub-(semi)lattice in $B(2^n)$.

- The essence of our conjecture is the question:

  *Does the **threshold behavior** of monotone properties in L(n) scale like monotone properties in B(n) or in $B(2^n)$?*

- Let F be a family of k-dimensional linear subspaces of $\{0,1\}^n$ such that F has density ≥ 1/2.
- The ***shadow*** ∂F is the set of k−1 dimensional subspaces of elements of F.

Main Case of Conjecture: Prove ∂F has density ≥ 0.51.

- The best known "q-analogue" of the Kruskal-Katona Theorem [Chowdhury & Patkos 2010] only shows that ∂F has density $(1/2)^{1-\Omega(1/k)}$.

THANK YOU