Learnability and positive equivalence relations

David Belanger, Ghent University

Joint work with Z Gao, S Jain, W Li and F Stephan.

I and II are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$

I and **II** are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$!

I and II are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$!
I: $\ldots 11, 13, 17, \ldots$

I and **II** are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$!
I: $\ldots 11, 13, 17, \ldots$
II: It's the prime numbers!

I and II are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$!
I: $\ldots 11, 13, 17, \ldots$
II: It's the prime numbers!
I: $\ldots 9, 15, 19, 21, \ldots$

I and II are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$!
I: $\ldots 11, 13, 17, \ldots$
II: It's the prime numbers!
I: $\ldots 9, 15, 19, 21, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$ after all!

I and **II** are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$!
I: $\ldots 11, 13, 17, \ldots$
II: It's the prime numbers!
I: $\ldots 9, 15, 19, 21, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$ after all!
I: $\ldots 8 \ldots$

I and II are playing a guessing game

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$!
I: $\ldots 11, 13, 17, \ldots$
II: It's the prime numbers!
I: $\ldots 9, 15, 19, 21, \ldots$
II: It's $\{2n + 1 : n \geq 1\}$ after all!
I: $\ldots 8 \ldots$

I: I'm thinking of a set that equals either the odd numbers, or the prime numbers.
I: $3, 5, 7, \ldots$
I: $17, 23, 11, 101, \ldots$
I: $107, 31, 97, 79, \ldots$
I: $199, 139, 29, 2$—
II: It's the primes!

The problem of learnability (Gold 1967)

Definition

A class $\mathcal{C} \subseteq \mathcal{P}(\mathbb{N})$ of r.e. sets is *R-learnable* if in the game:
I: I'm thinking of a set $A \in \mathcal{C}$,
I: $F_1 \subseteq F_2 \subseteq \cdots$ with $\bigcup_n F_n = A$,
then Player **II** can produce a description of A while following the ruleset R.

This *description* is an index $e = M(F_n)$ for the r.e. set $W_e = A$.

The problem of learnability (Gold 1967)

Definition

A class $\mathcal{C} \subseteq \mathcal{P}(\mathbb{N})$ of r.e. sets is *R-learnable* if in the game:
I: I'm thinking of a set $A \in \mathcal{C}$,
I: $F_1 \subseteq F_2 \subseteq \cdots$ with $\bigcup_n F_n = A$,
then Player **II** can produce a description of A while following the ruleset R.

This *description* is an index $e = M(F_n)$ for the r.e. set $W_e = A$.

Example rulesets

Finite: Only one guess is allowed.
Explanatory: Infinitely many guesses are allowed, but must reach a limit e.
Behaviourally correct: Infinitely many guesses are allowed, but cofinitely many must be correct.

Some specific examples

Finite: Only one guess is allowed.
Confident: Infinitely many guesses are allowed, and you must reach a limit e for every $B \subseteq \mathbb{N}$.
Explanatory: Infinitely many guesses are allowed, but must reach a limit e.
Vacillatory: Infinitely many guesses are allowed, but cofinitely many must be correct, and only finitely many *distinct* guesses.
Behaviourally correct: Infinitely many guesses are allowed, but cofinitely many must be correct.

$\mathcal{C} = \text{odds} \cup \text{primes}$	$\mathcal{C} = \{F \subseteq \mathbb{N} : \lvert F \rvert \leq 2\}$	$\mathcal{C} = \{F \overset{\text{fin}}{\subseteq} \mathbb{N}\}$
Fin learnable	Not fin learnable	Not fin learnable
Conf learnable	Conf learnable	Not conf learnable
Expl learnable	Expl learnable	Expl learnable
Vac learnable	Vac learnable	Vac learnable
BC learnable	BC learnable	BC learnable

Some specific examples

Finite: Only one guess is allowed.
Confident: Infinitely many guesses are allowed, and you must reach a limit e for every $B \subseteq \mathbb{N}$.
Explanatory: Infinitely many guesses are allowed, but must reach a limit e.
Vacillatory: Infinitely many guesses are allowed, but cofinitely many must be correct, and only finitely many *distinct* guesses.
Behaviourally correct: Infinitely many guesses are allowed, but cofinitely many must be correct.

$\mathcal{C} = \{\langle n, i \rangle : i < \lvert W_n \rvert\} \cup \{\langle n, i \rangle : i \in \mathbb{N}\}$	$\mathcal{C} = \{F \overset{\text{fin}}{\subseteq} \mathbb{N}\} \cup \mathbb{N}$
Not fin learnable	Not fin learnable
Not conf learnable	Not conf learnable
Not expl learnable	Not expl learnable
Vac learnable	Not vac learnable
BC learnable	Not BC learnable

Our setting

A *positive equivalence relation* is an r.e. equivalence relation.

Motivating question

Let $\langle a, b, c : r_1, r_2, r_3 \rangle$ be a group presentation. Then the word problem $E(w_1, w_2) \iff w_1 = w_2$ is a positive equivalence relation.

We want to 'learn' in this setting, while respecting equivalence classes. We restrict our attention to learning *E-families*, which are uniformly one-one enumerable families of sets closed under E.

Do the same theorems hold? For example, is it always true that:

$$\mathrm{Fin} \subsetneq \mathrm{Conf} \subsetneq \mathrm{Ex} \subsetneq \mathrm{Vac} \subsetneq \mathrm{BC}?$$

Some results: B., Gao, Jain, Li, Stephan

We always have:

$$\mathrm{Fin} \subseteq \mathrm{Conf} \subseteq \mathrm{Ex} \subseteq \mathrm{Vac} \subseteq \mathrm{BC}?$$

Which notions of learnability are distinct in $\mathbb{N}/E$?

Theorem 1

In every E we have expl $\neq$ BC.

Theorem 2

There is an E in which expl $=$ vac.

Some lemmas, part 1

Given an equivalence relation E, let $[a] = \{b : aEb\}$. Recursively define

$$a_k = \min\{b : \forall i < k, \quad b \notin [a_i]\},$$

and let

$$A_n = [a_0] \cup \cdots \cup [A_{n-1}].$$

Lemma 1

Every uniformly r.e. superclass of $\{A_n : n \in \mathbb{N}\}$ which respects E is an E-family.

In particular, the class $\{\bigcup_{k \in F}[a_k] : F \overset{\text{fin}}{\subseteq} \mathbb{N}\}$ of 'E-finite' sets is an E-family.

Some lemmas, part 2

Explanatory: Infinitely many guesses are allowed, but must reach a limit e.

Lemma 2 (Blum-Blum 1975)

If a set $A \in \mathcal{C}$ is explanatory-learned by M, then $\exists F \overset{\text{fin}}{\subseteq} A$ such that

$$F \subseteq G \overset{\text{fin}}{\subseteq} A \text{ implies } M(F) = M(G) = A.$$

This also holds with positive equivalence relations.

Lemma 3

If the E-finite sets are expl-learnable then each $\overline{[a]}$ is r.e.

Proof

Choose an $F \subseteq [a]$ as in Lemma 2. Then

$$\overline{[a]} = \{b : \text{eventually } M([a]_s \cup [b]_s) \neq M(F)\},$$

where $[a]_s$ and $[b]_s$ are approximations to $[a], [b]$.

Proof of Theorem 1: expl $\neq$ BC.

If the E-finite sets are not expl-learnable, we are done. Otherwise, let $\mathcal{C}$ consist of:

1. $A_n = [a_0] \cup \cdots \cup [a_{n-1}]$;
2. $F_n = \overline{[a_n]} = \bigcup_{k \neq n} [a_k]$;
3. If $|W_n| = m$ is finite, then for all k: $B_{n,k} = F_n \cap A_k$.

This is an E-family by Lemmas 1 and 3.

Proof of Theorem 1: expl $\neq$ BC.

If the E-finite sets are not expl-learnable, we are done. Otherwise, let $\mathcal{C}$ consist of:

1. $A_n = [a_0] \cup \cdots \cup [a_{n-1}]$;
2. $F_n = \overline{[a_n]} = \bigcup_{k \neq n} [a_k]$;
3. If $|W_n| = m$ is finite, then for all k: $B_{n,k} = F_n \cap A_k$.

This is an E-family by Lemmas 1 and 3.

By Lemma 2, every F_n has an $F \overset{\text{fin}}{\subseteq} F_n$ such that $F \subseteq G \overset{\text{fin}}{\subseteq} F_n$ implies $M(F) = M(G) = e$, and this limit e can be found using $0'$.

B_n also has such a subset. So if $|W_n|$ is finite, F must contain some $[a_m]$ with $m \geq |W_n|$. Using $0'$ to check whether $|W_n|$ is greater or smaller than the maximum $[a_m]$ in F, we can then compute the Σ_2-complete set $\{e : |W_e| \text{ finite}\}$. Thus $0'' \leq_T 0'$, a contradiction. □

Proof of Theorem 2: $\exists E$ s.t. expl = vac.

Construction. Begin with $a_m = m$ for all m. For each $s = 0, 1, \ldots$ in turn, search for $n, k, \ell < s$ s.t.:

- $n < \ell$ and $k < \ell$;
- $[a_k] \cap W_n = \emptyset$ as approximated at stage s; and
- $[a_\ell] \cap W_n \neq \emptyset$ as approximated at stage s.

If n, k, ℓ exist, select a triple with ℓ as small as possible. Merge $[a_k]$ and $[a_\ell]$ together, renumbering other a_m as necessary. Repeat until no more n, k, ℓ exist. Then move on to the s.

Claim 1. If W_n is E-closed, then either $W_n \subseteq A_n$, or $W_n = A_m$ for some $m \geq n$, or $W_n = \mathbb{N}$.

Claim 2. If $\mathcal{C}$ is an E-family then there are infinitely many $B \in \mathcal{C}$ such that $A_n \subseteq B$.

Proof: Induction on n, with Claim 1 and the Pigeonhole Principle.

Claim 3. expl = vac

Proof: Suppose M vac-learns $\mathcal{C}$. By Claims 1 and 2, $\mathbb{N} \notin \mathcal{C}$. Given an $A \in \mathcal{C}$, M will output some largest index n^*. Watch for this n^*, and use Claim 1 to form guesses about A.

Thank you!

I: I'm thinking of a set $A \subseteq \mathbb{N}$.
I: $3, 5, 7, \dots$
II: It's $\{2n + 1 : n \geq 1\}$!
I: $\dots 11, 13, 17, \dots$
II: It's the prime numbers!
I: $\dots 9, 15, 19, 21, \dots$
II: It's $\{2n + 1 : n \geq 1\}$ after all!
I: $\dots 8 \dots$

I: I'm thinking of a set that equals either the odd numbers, or the prime numbers.
I: $3, 5, 7, \dots$
I: $17, 23, 11, 101, \dots$
I: $107, 31, 97, 79, \dots$
I: $199, 139, 29, 2$—
II: It's the primes!