Formal Learning Theory 1

1 The Gold Learning Paradigm

Traditional important questions in linguistics[1]:

1. Is language acquisition more like learning from evidence or more like growth or triggering?
2. Are specific properties of human syntax innately given at the outset in every human learner?
3. How do properties of the language learning process shape the language we speak?
4. How can linguistic ability be acquired as rapidly and reliably as it is, from readily available data?

These questions are super important for linguists, but they are quite vague.

- Can we have a more precise formulation of them (and defs. of terms like *learning*) which will allow us to sharpen them up and allow us to provide more exact answers?

The Gold Paradigm (Gold (1967) and Blum & Blum (1975)).

- First (most influential) paradigm for the formal study of language learning.
- The GP yielded a formal result (**Gold's theorem**) which was taken to have important significance for our traditional important questions.
  - Many cognitive scientists treat it as providing support for **rationalism:** substantial innate knowledge or constraints are needed to facilitate language acquisition. (vs **empiricism**).
  - Gold's result is also frequently misinterpreted in the literature (cf. Johnson 2004).
- Today we see the description of the paradigm and some positive results about learnability, and next week we see some negative results and discuss their implications for cognitive science.

Languages are sets of expressions together with a non-expression #, where an expression is an element of Σ^*.

A **(positive) text T** is an infinite sequence whose elements are expressions and #.

- An infinite sequence is a function with domain $\mathbb{N}$, assigns to each expression a natural number.
- T(n) is the nth member of T, counting up from 0.
- T[n] is the length n initial sequence of T.

(1) $\text{T[n]} = \langle T(0), T(1), \ldots T(n-1)\rangle$

[1] The presentation of the Gold paradigm and the Gold theorem are taken from Stabler (2014) (refs. on the website).

The **content of text T,** content(T) is the set of expressions that appear in T.

- T is a text for language L iff content(T) = L.
- # is never part of the content.
- For every text T and every $x \in content(T)$, there is some finite i such that $x = T(i)$.
- A **learner** is a (possibly partial) function ϕ from finite sequences of expressions and # to grammars. (recall L(G) = the language generated by G)

1.1 A def. of learning: Identification from positive text

If $x \in Dom(\phi)$, we say ϕ is defined on x (notation $\phi(x)\downarrow$).

(2) ϕ **converges** on T, $\phi(T)\downarrow$, iff for some i, $\phi(T[j]) = \phi(T[i])$ for all $j \geq i$.

•In this case, we define $\phi(T)$ to be $\phi(T[i])$.

(3) ϕ identifies T iff $\phi(T)\downarrow$ and content(T) = L($\phi(T)$).

(4) ϕ identifies L iff ϕ identifies all texts for L.

(5) ϕ identifies a class of languages $\mathcal{L}$ iff for all $L \in \mathcal{L}$, ϕ identifies L.

(6) A class of languages $\mathcal{L}$ is **identifiable/learnable** iff some learner identifies $\mathcal{L}$.

2 Identification by enumeration

2.1 Enumerations

An **enumeration** is a sequence, a list, a function whose domain is the positive integers or $\mathbb{N}$.

- An enumerable set is one whose members can be enumerated (and also Ø).
- For any vocabulary Σ, Σ^* is enumerable. (sort into groups based on length of string, and alphabetical order).

(7) **Fact:** The set of all grammars is enumerable.

1. Assume that the elements of Σ are alphabetic, and order them alphabetically.
2. V is finite and disjoint from Σ, and we can order it alphabetically (for example).
3. We enumerate the elements of $\Sigma \cup V$ by ordering Σ before V.
4. We can then order the left sides of the rewrite rules in a similar way $((\Sigma \cup V^* \times V \times (\Sigma \cup V^*)))$, as well as the right sides $(\Sigma \cup V)$.
5. Assuming that the categories V, vocabulary Σ and start symbol S are given, we can then enumerate grammars with 0 symbols, grammars with 1 symbol, grammars with 2 symbols, and so on, to get all grammars.

2.2 Fin is identifiable/learnable

Théorème 2.1 *The class FIN is identifiable (Gold 1967).*

Proof: Suppose that we have an enumeration of grammars for all the r.e. languages.

Define a learner ϕ_e as follows: for any text T, $\phi_e(T[i])$ is the first grammar G in the enumeration such that L(G) = content(T[i])

- Consider any $L \in FIN$ and any text T for L. Then there is some i such that content(T[i]) = L. In this case, $\phi_e(T[j]) = \phi_e(T[j])$, for all $j \geq i$, so ϕ_e converges on T, and L($\phi_e(T[i]) = content(T)$).

Notice also that nothing in the definition of learnability implies that learners are necessarily computable functions, and so we did not need to explain how the function ϕ_e could be computed (what is the algorithm for obtaining the grammar?). If we are interested in a computable learner, we could let $\phi_e(T[i])$ be the grammar $\{S \rightarrow |x \in content(T[i])\}$. (for example)

3 Some positive results

3.1 $\mathcal{L}_{co-1}$ is identifiable

(8) a. Let $\mathcal{L}_{co-1}$ be the class of languages $\{\Sigma^* - \{x\}|x \in \Sigma^*\}$
b. For any number $k \geq 0$, let $\mathcal{L}_{co-k}$ be the class of languages $\{\Sigma^* - L||L| = k\}$

A grammar for a language in $\mathcal{L}_{co-1}$ could be regarded as a "filter" represented by a rule of the form (where x is any string):

(9) $*x$

A grammar for any language in $\mathcal{L}_{co-k}$ in one of the classes of $\mathcal{L}_{co-k}$ could be given by exactly k filters of this kind.

Théorème 3.1 *$\mathcal{L}_{co-1}$ is identifiable.*

Proof: Consider any enumeration of Σ^* and let ϕ be defined as follows:

- For any sequence of strings $T[i]$,
$$\phi(T[i]) = x*$$
, where $*x$ is the first string in the enumeration that does not occur in $T[i]$.

Now we show that, given any text T for any language $L \in \mathcal{L}_{co-1}$, ϕ identifies T.

- Since $L \in \mathcal{L}_{co-1}$, L is defined by some filter $*x$. Since there are just finitely many elements that occur before x, there will be some finite point i such that T[i] contains all those elements.
- At that point, by the definition of ϕ, $\phi[T(i)] = *x$.
- Furthermore, it is easy to see that at all later points $j > i$, $\phi(T[j]) = *x$ because x will never occur in the text.

Generalization:

Théorème 3.2 *For any k, $\mathcal{L}_{co-k}$ is identifiable.*

Proof: One learner for $\mathcal{L}_{co-k}$ is the function that, at each point, conjectures that the first k elements of Σ^* not seen so far are the ones excluded.