Learnability

Greg Kobele

Summer Semester, 2024

1 Course Description

Greg Kobele	gkobele@uni-leipzig.de
Office: GWZ 1511	Office Hours: Tues 12:00-12:50
Classroom: NSG S.215	Tues 13:15-14:45
https://home.uni-leipzig.de/gkobele/courses/2024.SS/Learnability	

Table 1: Brief Outline of Course Topics

Gold Learning
PAC Learning
Regular languages
CG/MGs from trees
Subregularity
Distributional Learning

2 Identification (Gold learning)

1. languages are sets of expressions, expressions are elements of Σ^*, where Σ is a finite vocabulary
2. A **(positive) text** is an infinite sequence of expressions and/or **pauses**
   - a pause is #
   - $T(i)$ is the expression at point i in T
   - $T[i]$ is the prefix of T of length i
   - expressions occuring in a text T are denoted $\mathit{content}(T)$
     - $\mathit{content}(T) = \{e \mid \exists i.t_i = e \wedge e \neq \#\}$
3. A text T is **for** a language L iff $L = \mathit{content}(T)$
   - for every text T and $x \in \mathit{content}(T)$, there is some i s.t. $T(i) = x$
4. A **learner** is a (partial) function ϕ from *finite* sequences of expressions to grammars
5. Learner ϕ is **defined** on T iff $\phi(T[i])$ is defined for all i
6. Learner ϕ **converges** on T iff
   (a) ϕ is defined on T, and
   (b) for some i, $\phi(T[i]) = \phi(T[j])$ for all $j \geq i$

   In this case, we define $\phi(T)$ to be $\phi(T[i])$, and say ϕ **converges to** $\phi(T[i])$
7. Learner ϕ **identifies** a text T iff
   (a) ϕ converges on T
   (b) and $L(\phi(T)) = \mathit{content}(T)$
8. A learner identifies a language L iff it identifies *every* text for L
9. A learner identifies a *class of languages* $\mathcal{L}$ iff it identifies every $L \in \mathcal{L}$
   - A class of languages is identifiable iff some learner identifies it

2.1 First results

Theorem 1. *The classes* $\mathcal{L} = \emptyset$ *and* $\mathcal{L} = \{L\}$ *are identifiable (for every r.e. lang* L*).*

Theorem 2 (Gold, 1967)**.** *The class* $\mathcal{L}_{fin}$ *of all finite languages is identifiable.*

Proof. Assume some enumeration $G_0, G_1, G_2, \ldots$ of the grammars for the finite languages. Define $\phi_e(T[i])$ to be the first grammar G s.t. $L(G) = \mathit{content}(T[i])$.

We show that ϕ_e identifies $\mathcal{L}_{fin}$. Let $L \in \mathcal{L}_{fin}$ be arbitrary, and let T be a text for L. Because L is finite, there will be some point i at which $\mathit{content}(T[i]) = L$. Then $\mathit{content}(T[j]) = \mathit{content}(T[i])$ for all $j \geq i$. As ϕ_e is defined extensionally, it converges on T. As it converges to a grammar for L, it identifies T. As T was arbitrary, it identifies L. As L was arbitrary, it identifies $\mathcal{L}_{fin}$. □

10. This kind of learner is called an **identification-by-enumeration** learner.

11. We did not specify how to construct or more generally to interact with this enumeration.

    - one way: by explicitly manipulating grammars

    $$\phi(T[i]) := \{S \to w | w \in \mathit{content}(T[i])\}$$

      Here the enumeration is only *implicit*

    - A common slogan: *Choose the* ***simplest*** *grammar compatible with the data*

12. A learner ϕ is **self-monitoring** iff

    (a) on any text T, there is a unique i s.t. $\phi(T[i]) = \star$

    (b) on any text T, if $\phi(T[i]) = \star$, then for all $j > i$, $\phi(T[j]) = \phi(T[i+1])$

Theorem 3 (Freivald and Wiehagen, 1979)**.** *No self-monitoring learner identifies* $\mathcal{L}_{fin}$*.*

Proof. Assume for a contradition that self-monitoring ϕ did identify $\mathcal{L}_{fin}$. Let $L \in \mathcal{L}_{fin}$ be arbitrary, and consider any text T where *content*$(T) = L$. Let i be the unique point in T where $\phi(T[i]) = \star$. Then by assumption $L(\phi(T[i+1])) = L(\phi(T[j]))$ for all $j > i$. Now let $w \notin L$, and consider the text $T' = T[i+1]www\ldots$. Then $\phi(T[i+1]) = \phi(T'[i+1])$ and thus $\phi(T) = \phi(T')$ but *content*$(T) \neq$ *content*(T'), and so ϕ fails to identify either T or T'. □

2.2 Formal Properties of Identification

13. let *SEQ* := $\{T[n] \mid T$ is a text and $n \in \mathbb{N}\}$ be the set of *finite* sequences of expressions (texts are *infinite* sequences of expresions). For $s, t \in$ *SEQ*,

    - $s^\frown t$ is their concatenation
    - $s \subseteq t$ iff s is a prefix of t

14. $s \in$ *SEQ* is a **locking sequence** for learner ϕ and language L iff

    (a) *content*$(s) \subseteq L$

    (b) $L(\phi(s)) = L$, and

    (c) for all $s' \in$ *SEQ* with *content*$(s') \subseteq L$, $\phi(s^\frown s') = \phi(s)$

Theorem 4 (Blum and Blum, 1975)**.** *if ϕ identifies L, then there is a locking sequence for ϕ and L*

Proof. For a contradiction, assume the theorem were false. Then ϕ identifies L, but there is no locking sequence. Spelling this out:

> for every finite sequence $s \in$ *SEQ* with *content*$(s) \subseteq L$ and where $L(\phi(s)) = L$, there is another sequence $s' \in$ *SEQ* with *content*$(s') \subseteq L$, but where $\phi(s^\frown s') \neq \phi(s)$.

But this will allow us to construct a text S for L which ϕ does not identify, establishing the contradiction.

Let T be a text for L. We define a set of finite sequences $s_0, s_1, \ldots$, each of which will be an initial segment of the desired S.

stage 0 $s_0 = \epsilon$

stage $n+1$ Given s_n, there are two cases.

case 1 $L(\phi(s_n)) \neq L$

Then let $s_{n+1} = s_n^\frown T(n)$.

case 2 $L(\phi(s_n)) = L$

Then by our condition, there is some $s' \in SEQ$ with $content(s') \subseteq L$ with the property that $\phi(s_n^\frown s') \neq \phi(s_n)$. Let $s_{n+1} = s_n^\frown s'^\frown T(n)$.

Observe that for all i, s_i is a prefix of s_{i+1}. Furthermore, for each s_i, $content(s_i) \subseteq L$. We define the text $S = s_0(0)\ \mathrm{s}_1(1)\ \mathrm{s}_2(2) \cdots$ for L. This text S is essentially the upper bound of the finite sequences s_i. However, by design, ϕ does not converge on S. □

Locking sequences are ubiquitous.

Corollary 4.1. *Let ϕ identify L, and $t \in SEQ$ with $content(t) \subseteq L$. Then there is some $s \in SEQ$ such that $t^\frown s$ is a locking sequence for ϕ and L.*

Proof. We use the same construction as in the previous theorem, except that $s_0 = t$. □

Locking sequences allow us to show that certain classes are unidentifiable.

Theorem 5 (Gold 1967)**.** *No superfinite class of languages is identifiable.*

Proof. Let $\mathcal{L}$ contain all finite languages, and at least one infinite language, L_∞. Assume for a contradiction that *phi* identifies $\mathcal{L}$. Let s be a locking sequence for ϕ and L_∞. There is a text T for $content(s)$ which begins with s (for example, s repeated *ad infinitum*). But then by the locking sequence theorem, $L(\phi(T)) = L_\infty$, whence ϕ doesn't identify $content(s)$. □

The classes of languages which are identifiable have a certain topological property.

Theorem 6 (Angluin 1980, Subset Theorem)**.** *$\mathcal{L}$ is identifiable iff for every $L \in \mathcal{L}$ there is a finite $D_L \subseteq L$ such that for all other $L' \in \mathcal{L}$, if $D_L \subseteq L'$ then $L' \not\subset L$.*

Proof. Let $\mathcal{L}$ be given.

left-to-right Suppose ϕ identifies $\mathcal{L}$. For each $L \in \mathcal{L}$ choose a locking sequence s_L for ϕ on L. We will show that $D_L = content(s_L)$. Towards a contradiction, assume that there is some intervening $L' \in \mathcal{L}$ such that $content(s_L) \subseteq L' \subset L$. Let T be a text for this hypothesized L' beginning with s_L. Because s_L is a locking sequence for L, and because for every $j \geq |s_L|$, $T[j] = s_L\frown t_j$ for some $t_j \in SEQ$ with $content(t_j) \subseteq L$, ϕ must converge to L on T, which means that ϕ doesn't identify T which is for L', and thus not L', and thus not $\mathcal{L}$. A contradiction.

right-to-left Assume that for every $L \in \mathcal{L}$ there is a finite $D_L \subseteq L$ such that for all other $L' \in \mathcal{L}$, if $D_L \subseteq L'$ then $L' \not\subset L$. We must show that some learner identifies $\mathcal{L}$. Assume some enumeration of grammars. We define a learner ϕ as follows:

> For all $s \in SEQ$, $\phi(s)$ is the first grammar G in the enumeration for an $L \in \mathcal{L}$ with the property that
>
> $$D_L \subseteq content(s) \subseteq L$$
>
> Otherwise $\phi(s)$ is the first grammar in the enumeration.

Now we show that ϕ so defined actually identifies $\mathcal{L}$. Let $L \in \mathcal{L}$ be arbitrary, and let T be a text for L. Then for some n, $D_L \subseteq content(L[m]) \subseteq L$ for all $m \geq n$. Consider the first grammar G for L in the enumeration. Learner ϕ will not hypothesize G on $T[k]$ for $k \geq n$ only if there is some earlier G' in the enumeration with $D_{L(G')} \subseteq T[k] \subseteq L(G')$. Assume such a G' exists. We show that at some point this G' no longer meets this condition, and thus ϕ must abandon it (and move closer to G). By hypothesis, $D_{L(G')} \subseteq T[k] \subseteq L$, whence by assumption of the theorem $L \not\subseteq L(G')$. Thus there is some $u \in L - L(G')$. As T is a text for L, u appears in T at some point j, at which point it no longer holds that $D_{L(G')} \subseteq T[j] \subseteq L(G')$ and so ϕ abandons its conjecture.

□