Learning Regular Sets

Author: Dana Angluin
Presented by: M. Andreína Francisco

Department of Computer Science
Uppsala University

February 3, 2014

Minimally Adequate Teachers

A Minimally Adequate Teacher (MAT) is an Oracle that must answer correctly two types of questions:

- Membership queries
  - the answer must be **yes** or **no**
- Strong equivalence queries
  - the answer is **yes** or any **counterexample**

Definitions

Prefix-closed

Every prefix of every member is also a member, that is

$$uv \in S \Rightarrow u \in S$$

i.e. the set $\{\lambda, \mathrm{a}, \mathrm{ab}\}$

Suffix-closed

Every suffix of every member is also a member, that is

$$uv \in S \Rightarrow v \in S$$

i.e. the set $\{\lambda, \mathrm{a}, \mathrm{ba}\}$

Observation Table

An observation table is a triple $\langle \mathrm{STA}, \mathrm{EXP}, \mathrm{OT} \rangle$ where:

- $\mathrm{STA} = \mathrm{RED} \cup \mathrm{BLUE}$
  - $\mathrm{RED} \subset \Sigma^*$ is a finite set of states
  - $\mathrm{BLUE} = \mathrm{RED} \cdot \Sigma \setminus \mathrm{RED}$ is the set of successor states of RED that are not RED
- $\mathrm{EXP} \subset \Sigma^*$ is the experiment set.
- $\mathrm{OT} : \mathrm{STA} \times \mathrm{EXP} \to \{0, 1, *\}$ is a function such that:

$$\mathrm{OT}[u][e] = \begin{cases} 1 & \text{if } ue \in L \\ 0 & \text{if } ue \notin L \\ * & \text{otherwise (not known)} \end{cases}$$

Observation Table

Examples

Experiment set

States

	λ	a
λ	0	1
a	1	0
b	1	0
aa	0	1
ab	1	0

How to build an automaton from an OT

- $Q \leftarrow \{q_r : r \in \mathrm{RED}\}$
- $\mathbb{F}_{\mathbb{A}} \leftarrow \{q_{we} : we \in \mathrm{RED} \wedge \mathrm{OT}[w][e] = 1\}$
- $\mathbb{F}_{\mathbb{R}} \leftarrow \{q_{we} : we \in \mathrm{RED} \wedge \mathrm{OT}[w][e] = 0$
- $\forall q_w \in Q \wedge \forall \sigma \in \Sigma \mid \delta(q_w, \sigma) \leftarrow q_u : u \in \mathrm{RED} \wedge \mathrm{OT}[u] = \mathrm{OT}[w\sigma]$

How to build an automaton from an OT

- $\mathbb{F}_{\mathbb{A}} \leftarrow \{q_{we} : we \in \mathrm{RED} \wedge \mathrm{OT}[w][e] = 1\}$
- $\forall q_w \in Q \wedge \forall \sigma \in \Sigma \mid \delta(q_w, \sigma) \leftarrow q_u : u \in \mathrm{RED} \wedge \mathrm{OT}[u] = \mathrm{OT}[w\sigma]$

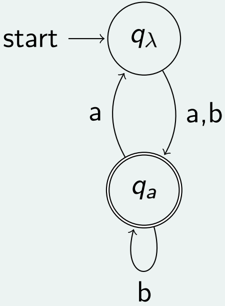
Examples

Observation Table

	λ	a
λ	0	1
a	1	0
b	1	0
aa	0	1
ab	1	0

Transition Table

	a	b
q_λ	q_a	q_a
q_a	q_λ	q_a



Closed Table

A table is closed if for every row $u \in \mathrm{BLUE}$ there is a row $v \in \mathrm{RED}$ such that $\mathrm{row}(u) = \mathrm{row}(v)$

Examples

The following table is not closed because of row 'ab'

	λ	a
λ	0	1
a	1	0
b	1	0
aa	0	1
ab	1	1

How do we close a table?

Let s be the row of BLUE that is not in RED

- Move s to RED
- $\forall a \in \Sigma$, add sa to BLUE

Examples

The following table is now closed but it is not complete.

	λ	a
λ	0	1
a	1	0
ab	1	1
b	1	0
aa	0	1
aba	*	*
abb	*	*

Consistent Table

A table is consistent if every pair of equivalent rows in RED remains equivalent after appending any symbol.

Examples

The following table is inconsistent

	λ	a
λ	0	1
a	1	0
ab	1	0
b	1	0
aa	0	1
aba	0	0
abb	1	0

Consistent Table

A table is consistent if every pair of equivalent rows in RED remains equivalent after appending any symbol.

Examples

The following table is inconsistent

	λ	a
λ	0	1
a	1	0
ab	1	0
b	1	0
aa	0	1
aba	0	0
abb	1	0

Consistent Table

A table is consistent if every pair of equivalent rows in RED remains equivalent after appending any symbol.

Examples

The following table is inconsistent

	λ	a
λ	0	1
a	1	0
ab	1	0
b	1	0
aa	0	1
aba	0	0
abb	1	0

How do we make a table consistent?

- Let $x \in \Sigma$ be the differentiating string such that $\mathrm{OT}[s_1] = \mathrm{OT}[s_2]$ but $\mathrm{OT}[s_1 x] \neq \mathrm{OT}[s_2 x]$.
- Let e be the experiment for which $\mathrm{OT}[s_1 x][e] \neq \mathrm{OT}[s_2 x][e]$.
- Adding the experiment 'xe' will differentiate $\mathrm{OT}[s_1]$ and $\mathrm{OT}[s_2]$.

Examples

	λ	a	aa
λ	0	1	0
a	1	0	*
ab	1	0	*
b	1	0	*
aa	0	1	*
aba	0	0	*
abb	1	0	*

The Learner $L*$

1. Initialize $\text{RED} = \text{EXP} = \{\lambda\}$
2. Initialize $\text{BLUE} = \text{RED} \cdot \Sigma$
3. Complete the table
4. while Teacher says NO
   - 4.1 while table is not closed or not consistent
     - 4.1.1 if not consistent $\rightarrow$ make it consistent (and complete)
     - 4.1.2 if not closed $\rightarrow$ close it (and complete it)
   - 4.2 Make a guess!
   - 4.3 If the teacher says no, then
     - 4.3.1 add the counterexample and all its prefixes to RED
     - 4.3.2 update $\text{BLUE} = \text{RED} \cdot \Sigma$
     - 4.3.3 complete the table

An Example Run of L^*

Suppose the unknown regular set is the set of all strings over $\{a, b\}$ with an even number of a's and an even number of b's

Initial table

	λ
λ	1
a	0
b	0

An Example Run of L^*

Suppose the unknown regular set is the set of all strings over $\{a, b\}$ with an even number of a's and an even number of b's

Initial table

	λ
λ	1
a	0
b	0

This table is not closed

An Example Run of L^*

We close the table

	λ
λ	1
a	0
b	0
aa	*
ab	*

An Example Run of L^*

We complete the table

	λ
λ	1
a	0
b	0
aa	1
ab	0

An Example Run of L^*

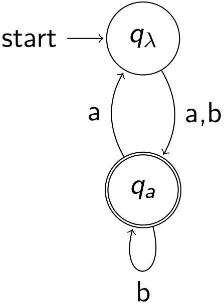
We complete the table

	λ
λ	1
a	0
b	0
aa	1
ab	0

This table is closed, consistent and complete, so we can make a guess

An Example Run of L^*

The guess is...



(note that it is the same automaton as in the construction example, but the observation tables are different)

An Example Run of L^*

Assume we get the counter example '*bb*'.

1. add the counterexample and all its prefixes to RED
2. update $\mathrm{BLUE} = \mathrm{RED} \cdot \Sigma$

	λ
λ	1
a	0
b	0
bb	*
aa	1
ab	0
ba	*
bba	*
bbb	*

An Example Run of L^*

Assume we get the counter example '*bb*'.

1. add the counterexample and all its prefixes to RED
2. update $\mathrm{BLUE} = \mathrm{RED} \cdot \Sigma$

	λ
λ	1
a	0
b	0
bb	1
aa	1
ab	0
ba	0
bba	0
bbb	0

An Example Run of L^*

Assume we get the counter example '*bb*'.

1. add the counterexample and all its prefixes to RED
2. update $\mathrm{BLUE} = \mathrm{RED} \cdot \Sigma$

	λ
λ	1
a	0
b	0
bb	1
aa	1
ab	0
ba	0
bba	0
bbb	0

An Example Run of L^*

Add the experiment '*a*'

	λ	a
λ	1	0
a	0	1
b	0	0
bb	1	0
aa	1	*
ab	0	*
ba	0	*
bba	0	*
bbb	0	*

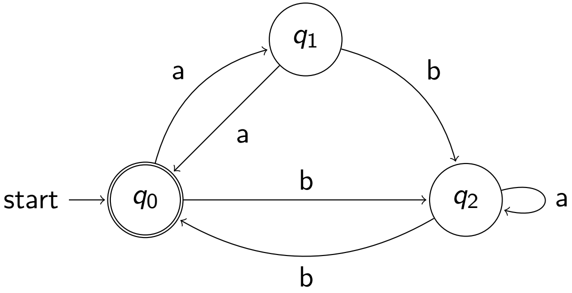
An Example Run of L^*

Add the experiment '*a*'

	λ	a
λ	1	0
a	0	1
b	0	0
bb	1	0
aa	1	0
ab	0	0
ba	0	0
bba	0	1
bbb	0	0

An Example Run of L^*

make a guess...



An Example Run of L^*

Suppose we get the counterexample '*abb*'

	λ	a
λ	1	0
a	0	1
b	0	0
bb	1	0
ab	0	0
abb	*	*
aa	1	0
ba	0	0
bba	0	1
bbb	0	0
aba	*	*
abba	*	*
abbb	*	*

An Example Run of L^*

Suppose we get the counterexample '*abb*'

	λ	a
λ	1	0
a	0	1
b	0	0
bb	1	0
ab	0	0
abb	0	1
aa	1	0
ba	0	0
bba	0	1
bbb	0	0
aba	0	0
abba	1	0
abbb	0	0

An Example Run of L^*

The table is not consistent

	λ	a
λ	1	0
a	0	1
b	0	0
bb	1	0
ab	0	0
abb	0	1
aa	1	0
ba	0	0
bba	0	1
bbb	0	0
aba	0	0
abba	1	0
abbb	0	0

An Example Run of L^*

The table is not consistent

	λ	a	b
λ	1	0	0
a	0	1	0
b	0	0	1
bb	1	0	0
ab	0	0	0
abb	0	1	0
aa	1	0	*
ba	0	0	*
bba	0	1	*
bbb	0	0	*
aba	0	0	*
abba	1	0	*
abbb	0	0	*

An Example Run of L^*

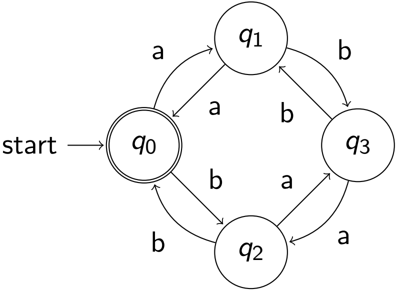
We complete the table

	λ	a	b
λ	1	0	0
a	0	1	0
b	0	0	1
bb	1	0	0
ab	0	0	0
abb	0	1	0
aa	1	0	0
ba	0	0	0
bba	0	1	0
bbb	0	0	1
aba	0	0	1
abba	1	0	0
abbb	0	0	0

An Example Run of L^*

We make a guess... and the Teacher says YES! :-)

δ	a	b
q_0	q_1	q_2
q_1	q_0	q_3
q_2	q_3	q_0
q_3	q_2	q_1



Questions?