Decideability

Sipser pages 165 - 186
Decideability

- A class of problems is decidable if every problem can be answered Yes or No.
- We often look at classes that ask questions about languages and automata.
- We generally use the notation \(<P>\) to describe an encoding of a problem P in some way as input to a Turing machine.
- Turing machines are a good mechanism to talk about decideability.
  - Why?
  - What characteristics do Turing machines have?
Problems about DFA’s

• $A_{DFA}$ Does a DFA (B) accept some string (w)?

• $E_{DFA}$ Is the language accepted by some DFA (B) the empty language (the empty set of strings).

• $EQ_{DFA}$ Do two DFAs (A and B) accept the same language.
• Does a DFA (B) accept some string (w)?

• $A_{DFA} = \{ <B,w> \mid B \text{ is a DFA that accepts input string } w \}$

• Note that $A_{DFA}$ is a language problem itself.
• Consider $<B,w>$ to be the input language
• And the solution a Turing Machine that halts on all such input in either the accept or reject state
Representations

• Recall \(<B, w>\) is meant to represent a DFA and some input string.
• How might we represent this as input to a TM?
• \(B = (Q, \Sigma, \delta, q_0, F)\)
Checking $A_{DFA}$

- Let $M$ be a TM that does the following
  - Does the input represent a legal DFA
    - If not then reject
  - Simulate $B$ on input
  - When finishing processing $w$, if the simulation is in an accepting state, then the TM accepts, else the TM rejects.
\[ A_{\text{NFA}} \]

• How might we show that an NFA (C) accepts a string \( w \)?
• We might use a similar approach, encoding the NFA and its input on a TM tape \( <C,w> \) and then simulating the NFA.
• There is another approach!
• Since every NFA has an equivalent DFA (the subset construction) let's use the TM (M) of the last section.
N an TM that decides $A_{\text{NFA}}$

- N is a TM on input $<C,w>$ where C is an NFA and w is a string
  1. Convert C into a DFA (D) using subset construction
  2. Run M on $<D,w>$
  3. If M accepts, then N accepts, if M rejects, then N rejects

This is an example of an important strategy called reduction
How might we decide $A_{RegExp}$
$E_{\text{DFA}}$ is decidable

• T = On input $<A>$ where $A$ is a DFA
  – Mark the start state of $A$
  – Repeat until no new state is marked
    • Mark any state that has a transition coming into it from any state that is already marked
  – If no final state of $A$ is marked, accept; otherwise reject
EQ_{DFA} is decidable

- To test if two DFAs decide the same language we will rely on several facts
  - DFA’s are closed under intersection, union, and complement
  - EDFA is decidable (TM T from previous section)
Symmetric Difference

\[ L(C) = (L(A) \cap L(B)) \cup (L(A) \cap L(B)) \]

If A and B are equal, then the symmetric difference is empty.
\( \text{EQ}_{\text{DFA}} \) is decidable

- \( F = \) On input \(<A, B>\) where \( A \) and \( B \) are DFAs
  1. Construct DFA \( C \), the symmetric difference of \( A \) and \( B \)
  2. Run TM \( T \) (the one that decides \( E_{\text{DFA}} \)) on \(<C>\)
  3. If \( T \) accepts, then \( F \) accepts, if \( T \) rejects, then \( F \) rejects
Problems about CFG’s

The following class of problems are discussed in the text. Be sure and read about them.

– $A_{\text{CFG}}$ Does a CFG (B) accept some string (w)?
– $E_{\text{CFG}}$ Is the language accepted by some CFG (B) the empty language (the empty set of strings).
  • This one is quite interesting, and not what one might expect. Pay close attention!
– $EQ_{\text{CFG}}$ Do two CFGs (A and B) accept the same language.
The size of infinite sets

• How can we tell if two sets have the same size?
• Easy for finite sets.
• Not so straightforward for infinite sets

Two infinite sets have the same size if every element of one can be paired with the elements of the other.
Properties of functions

• One-to-one
  – A function, f, is one-to-one if it never maps two different elements of the domain to the same element of the range. \( x \neq y \Rightarrow f(x) \neq f(y) \)

• Onto
  – A function f is onto, if every element of the range is mapped to by some element of the domain

• Correspondence
  – A function is a correspondence if it is both one-to-one and onto
Naturals and the even-Naturals have the same size

<table>
<thead>
<tr>
<th>n</th>
<th>f(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

\[ f(n) = n \times 2 \]

F is one-to-one, two numbers never map to the same element.

F is onto, every even number is mapped to.
Countable sets

• Definition
  – A set is countable, if it is finite, or if it is infinite, it is in correspondence to the Natural numbers
Rational numbers are countable

- Rational numbers, numbers exactly expressed as $x/y$, are countable.

How can we establish a correspondance?

Can’t travel along one row.

Or along one column.

But along the diagonals.
The real numbers are not countable

- We show no correspondence between R and N can exist.
- We use a classic argument (due to Cantor) called a diagonalization argument.

- First recall that every Real number can be expressed as an infinite decimal expansion. Example
  - 3.1415962...
  - 2.0000000...
  - 0.1250000...
  - 5.5555555...
Proof by contradiction.

• Assume that the Naturals and the Reals are in correspondence, then there exists a one-to-one, onto function, \( f : \text{Nat} \rightarrow \text{Real} \)

A part of the correspondence, \( f \), between the Naturals and the Reals

<table>
<thead>
<tr>
<th>( n )</th>
<th>( f(n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.14159...</td>
</tr>
<tr>
<td>2</td>
<td>55.5555...</td>
</tr>
<tr>
<td>3</td>
<td>0.12500...</td>
</tr>
<tr>
<td>4</td>
<td>0.50000...</td>
</tr>
</tbody>
</table>

We show that \( f \) can’t be onto, thus it can’t be a correspondence, and hence the Reals can’t be countable
Consider the real between 0 and 1

- All its digits are after the decimal point
- The nth digit after the decimal point is chosen different from the nth digit of the nth number, for example .2669...
  - 2 ≠ 1
  - 6 ≠ 5
  - 9 ≠ 0
- Note that no natural maps to this number. Suppose one did, let it be Z, but the Zth digit of f(Z) differs from our number in the Zth digit by construction.
- This is a contradiction, so our assumption that the Reals are countable must be false.
The set of all Turing Machines is countable

• Recall if $\Sigma$ is finite, then $\Sigma^*$ is countable
• We can write them all down
  – First all of length 0
  – Then all of length 1
  – Then all of length 2
  – Then all of length 3
• Each Turing Machine (M) has an encoding as $<M>$ which is a string in $\Sigma^*$
The set of all infinite binary strings is not countable.

- Diagonalization argument
- Consider 0101...
- Differs from the nth digit in the nth string

<table>
<thead>
<tr>
<th>n</th>
<th>f(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1011001...</td>
</tr>
<tr>
<td>2</td>
<td>0010100...</td>
</tr>
<tr>
<td>3</td>
<td>1010111...</td>
</tr>
<tr>
<td>4</td>
<td>0110110...</td>
</tr>
</tbody>
</table>
Consider the following function: $F$

Given a finite alphabet $\Sigma$

Given a language $L$ over $\Sigma$

- $L \subseteq \Sigma^*$
- $\Sigma^*$ is countable (thus so is $L$)

$F(i) = 1$ if the $i$th string of $\Sigma^*$ is in $L$, and $0$ otherwise.

We call $F$ the characteristic function of $L$
The set of languages is not countable

• Given a finite alphabet $\Sigma$

• Consider the set of all languages, $\mathcal{L}$, over $\Sigma^*$

• Each language $L$ in $\mathcal{L}$ has a characteristic function, $F$, which is an infinite sequence of 0’s and 1’s (i.e. an infinite binary sequence)
  – Eg consider $L = \{ x \mid \text{length of } x \text{ is even} \}$
  – $F(\varepsilon)=1;\ F(0)=0;\ F(1)=0;\ F(11)=1;\ F(00)=1;\ F(01)=1;\ F(10)=1;\ ...$

• Thus, there is a correspondence between languages and infinite binary sequences.

• We know that the set of infinite binary sequences is not countable, so the set of languages over a finite alphabet $\Sigma^*$, can’t be countable either!
There are languages not accepted by a Turing Machine. Sipser pg 178

• There are countable number of TMs

• A Turing Machine describes a language.

• There are uncountable number of languages.

• Thus some languages must not be describable by a TM.
The Halting problem

• Until now every problem we have looked at closely has been decidable.

• One might ask: “is any problem undecidable?”

• There is at least 1 undecidable problem $A_{TM}$
  – Acceptance by Turing Machine
  – Does an arbitrary TM accept an arbitrary input is undecidable

• This is an important result, both philosophically and computationally!
$A_{TM}$ is Turing Recognizable!

- While not decidable, $A_{TM}$ is Turing Recognizable.
- This depends upon the fact that there is a universal TM
- The universal Turing Machine takes $<tm,input>$ and simulates “tm” on “input”.
- Note if “tm” does not halt on “input” neither does the universal TM halt on $<tm,input>$
$R_{TM}$, Recognizing a TM

- U = On input $<M, w>$, where M is a TM and w is a string
  - Simulate M on input w
  - If M ever enters its accept state, accept; if M ever enters its reject state, reject

- Note, if we had a way of determining that M would not halt on w, we could reject, but we don’t.
$A_{TM}$ is undecidable
Sipser pg 179

• Proof by contradiction
• Assume that $A_{TM}$ is decidable. By a TM called $H$
  \[ H(<M,w>) = \text{accept if } M \text{ accepts } w, \text{ and reject if } M \text{ does not accept } w \text{ (i.e. } M \text{ either rejects or loops) } \]

• Then if $M$ decides, we can make another machine $D$
  • $D(<M>) = \text{ accept if } H(<M>,<M>) \text{ rejects, and rejects if } H(<M>,<M>) \text{ accepts}$
How a Turing machine $M$ and $H(<M>,w)$ are related.

<table>
<thead>
<tr>
<th>$M(w)$</th>
<th>accept</th>
<th>reject</th>
<th>loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(&lt;M,w&gt;)$</td>
<td>accept</td>
<td>reject</td>
<td>reject</td>
</tr>
</tbody>
</table>

How a Turing machine $M$ and $H(<M>,<M>)$ and $D(<M>)$ are related.

<table>
<thead>
<tr>
<th>$M(&lt;M&gt;)$</th>
<th>accept</th>
<th>reject</th>
<th>loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(&lt;M&gt;,&lt;M&gt;)$</td>
<td>accept</td>
<td>reject</td>
<td>reject</td>
</tr>
<tr>
<td>$D(&lt;M&gt;)$</td>
<td>reject</td>
<td>accept</td>
<td>accept</td>
</tr>
</tbody>
</table>

The curious case when $D$ is applied to itself.

<table>
<thead>
<tr>
<th>$D(&lt;D&gt;)$</th>
<th>accept</th>
<th>reject</th>
<th>loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(&lt;D&gt;,&lt;D&gt;)$</td>
<td>accept</td>
<td>reject</td>
<td>reject</td>
</tr>
<tr>
<td>$D(&lt;D&gt;)$</td>
<td>reject</td>
<td>accept</td>
<td>accept</td>
</tr>
</tbody>
</table>
Conclusion: $A_{TM}$ is undecidable

<table>
<thead>
<tr>
<th></th>
<th>Accept</th>
<th>Reject</th>
<th>Loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D(&lt;D&gt;)$</td>
<td>accept</td>
<td>reject</td>
<td>loop</td>
</tr>
<tr>
<td>$H(&lt;D&gt;,&lt;D&gt;)$</td>
<td>accept</td>
<td>reject</td>
<td>reject</td>
</tr>
<tr>
<td>$D(&lt;D&gt;)$</td>
<td>reject</td>
<td>accept</td>
<td>accept</td>
</tr>
</tbody>
</table>

- Since $D(<D>)$ rejects if $D(<D>)$ accepts we have reached a contradiction.
- So our original assumption that $A_{TM}$ is decidable must be incorrect.
- Thus, $A_{TM}$ is must be undecidable
Visualizing Diagonalization of $A_{TM}$

A table of the results of applying $H(<M_i><M_j>)$

<table>
<thead>
<tr>
<th>$H(I,j)$</th>
<th>$&lt;M1&gt;$</th>
<th>$&lt;M2&gt;$</th>
<th>$&lt;M3&gt;$</th>
<th>$&lt;M4&gt;$</th>
<th>$&lt;M5&gt;$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Accept</td>
<td>Reject</td>
<td>Reject</td>
<td>Accept</td>
<td>Reject</td>
</tr>
<tr>
<td>M2</td>
<td>Reject</td>
<td>Accept</td>
<td>Accept</td>
<td>Reject</td>
<td>Accept</td>
</tr>
<tr>
<td>M3</td>
<td>Reject</td>
<td>Reject</td>
<td>Reject</td>
<td>Reject</td>
<td>Reject</td>
</tr>
<tr>
<td>M4</td>
<td>Accept</td>
<td>Accept</td>
<td>Reject</td>
<td>Accept</td>
<td>Accept</td>
</tr>
<tr>
<td>M5</td>
<td>Reject</td>
<td>Accept</td>
<td>Reject</td>
<td>Accept</td>
<td>Reject</td>
</tr>
</tbody>
</table>
D is a TM so where is it in the Table?

<table>
<thead>
<tr>
<th>$H(&lt;M&gt;,&lt;M&gt;)$</th>
<th>accept</th>
<th>reject</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D(&lt;M&gt;)$</td>
<td>reject</td>
<td>accept</td>
<td>accept</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H(i,j)$</th>
<th>$&lt;M_1&gt;$</th>
<th>...</th>
<th>$&lt;D&gt;$</th>
<th>...</th>
<th>$&lt;M_5&gt;$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Accept</td>
<td>Reject</td>
<td>Reject</td>
<td>Accept</td>
<td>Reject</td>
</tr>
<tr>
<td>...</td>
<td>Reject</td>
<td>Accept</td>
<td>Accept</td>
<td>Reject</td>
<td>Accept</td>
</tr>
<tr>
<td>D</td>
<td>Reject</td>
<td>Reject</td>
<td>?</td>
<td>Reject</td>
<td>Reject</td>
</tr>
<tr>
<td>...</td>
<td>Accept</td>
<td>Accept</td>
<td>Reject</td>
<td>Accept</td>
<td>Accept</td>
</tr>
<tr>
<td>M5</td>
<td>Reject</td>
<td>Accept</td>
<td>Reject</td>
<td>Accept</td>
<td>Reject</td>
</tr>
</tbody>
</table>
Definition

• A language is Turing co-recognizable if its complement is Turing recognizable.

• Recall the complement of a language is the language with all the strings not recognized by the original language.

<table>
<thead>
<tr>
<th>M(w)</th>
<th>accept</th>
<th>reject</th>
<th>loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>CompM(w)</td>
<td>reject</td>
<td>accept</td>
<td>accept</td>
</tr>
</tbody>
</table>
Lemma

• A language, L, is decidable if and only if it is both Turing recognizable and Co-Turing recognizable.

• Two things to prove
  1. If L is decidable then it is both Turing and Co-Turing recognizable. This way is easy
  2. If L is Turing and Co-Turing recognizable, it is decidable
If $M$ is Turing and Co-Turing recognizable, it is decidable

<table>
<thead>
<tr>
<th>Turing recognizer</th>
<th>accept</th>
<th>reject</th>
<th>loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1(w)$</td>
<td>accept</td>
<td>reject</td>
<td>loop</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Turing Co-recognizer</th>
<th>accept</th>
<th>reject</th>
<th>loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_2(w)$</td>
<td>accept</td>
<td>reject</td>
<td>loop</td>
</tr>
</tbody>
</table>

- $P(w) = \text{run } M_1(w) \text{ and } M_2(w) \text{ in parallel}$
- If $M_1$ accepts, then $P$ accepts.
- If $M_2$ accepts then $P$ rejects.
P is a decider

- \( P(w) = \) run \( M_1(w) \) and \( M_2(w) \) in parallel
- If \( M_1 \) accepts, then \( P \) accepts.
- If \( M_2 \) accepts then \( P \) rejects.

- Every string, \( w \), is either in \( L \) (\( M_1 \) halts and accepts) or it is not (\( M_2 \) halts and rejects)
- So one of \( M_1(w) \) or \( M_2(w) \) must halt.
- \( P \) halts when either \( M_1 \) or \( M_2 \) halts, so \( P \) must Halt.
- So \( P \) is a decider that accepts all strings in \( L \) and rejects all strings not in \( L \)
Some languages aren’t even recognizable!

Sipser pg 81

• Consider the language which is the complement of $\overline{A_{TM}}$ which we write $\overline{A_{TM}}$

• We prove that $\overline{A_{TM}}$ is not Turing recognizable using a proof by contradiction

•
Proof

• Assume that $A_{TM}$ is Turing recognizable
• We know $A_{TM}$ is Turing recognizable
  – Sipser pg 174, theorem 4.11, Slide 29 in these notes
• Thus by our lemma $A_{TM}$ is decidable
• We know that ATM is not decidable, which leads to a contradiction
• So our original assumption that $A_{TM}$ is Turing recognizable must be flawed.
Review: Positive results

• Countable and uncountable Sets.
• Acceptance of Regular and Context Free languages is decidable.
• Equality of Regular and Context Free languages is decidable.
• Emptiness of Regular and Context Free languages is decidable.
Review: Negative results

• There are uncountable Sets
  – The reals, infinite binary sequences, languages over a finite alphabet.

• There are languages not described by any Turing Machine.

• There is an un-decidable language
  – $A_{TM}$ is undecidable
    – But, $A_{TM}$ is Turing recognizable

• There is a language that is not even Turing recognizable! ($\overline{A_{TM}}$ the complement of $A_{TM}$)