

Tugas Besar IF2124 Teori Bahasa Formal dan Otomata
HTML Checker dengan Pushdown Automata (PDA)
Semester I Tahun 2023/2024



KELOMPOK : 4 - pinjamdulu100

Hugo Sabam Augusto Sianturi 13522129

Muhammad Zaki 13522136

Andi Marihot Sitorus 13522138

SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA INSTITUT TEKNOLOGI
BANDUNG TAHUN AJARAN 2023/2024

BAB I

DESKRIPSI MASALAH

HTML (Hypertext Markup Language) adalah bahasa markup yang digunakan untuk membuat struktur dan tampilan konten web. HTML adalah salah satu bahasa utama yang digunakan dalam pengembangan web dan digunakan untuk menggambarkan bagaimana elemen-elemen konten, seperti teks, gambar, tautan, dan media, akan ditampilkan di browser web. Setiap dokumen HTML dimulai dengan elemen `<html>`, lalu diikuti dengan `<head>` (untuk metadata dan tautan ke file eksternal) dan `<body>` (untuk konten yang akan ditampilkan)

HTML menggunakan elemen-elemen (*tags*) untuk mengelompokkan dan mengatur konten. Contohnya, `<p>` digunakan untuk paragraf teks, `<h1>` hingga `<h6>` digunakan untuk judul, `<a>` untuk tautan, `` untuk gambar, dan sebagainya. Elemen HTML sering memiliki atribut yang memberikan informasi tambahan tentang elemen tersebut. Contohnya adalah atribut `src` untuk gambar, `href` untuk tautan, dan `class` untuk memberikan elemen kelas CSS.

Sama seperti bahasa pada umumnya, HTML juga memiliki sintaks tersendiri dalam penulisannya yang dapat menimbulkan error jika tidak dipenuhi. Meskipun web browser modern seperti Chrome dan Firefox cenderung tidak menghiraukan error pada HTML memastikan bahwa HTML benar dan terbentuk dengan baik masih penting untuk beberapa alasan seperti *Search Engine Optimization (SEO)*, aksesibilitas, *maintenance* yang lebih baik, kecepatan render, dan profesionalisme.

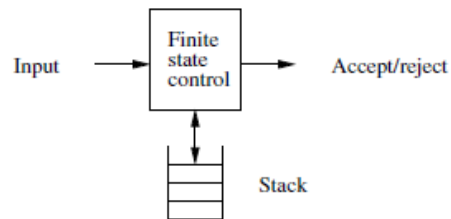
Dibutuhkan sebuah program pendeteksi error untuk HTML. Oleh sebab itu, implementasikan sebuah program yang dapat memeriksa kebenaran HTML dari segi nama tag yang digunakan serta atribut yang dimilikinya. Pada tugas pemrograman ini, gunakanlah konsep Pushdown Automata (PDA) dalam mencapai hal tersebut yang diimplementasikan dalam bahasa Python.

BAB II

TEORI DASAR

Pushdown automaton pada dasarnya adalah mesin otomata tak hingga non deterministik dengan transisi yang diperbolehkan dan satu kemampuan tambahan: sebuah tumpukan (stack) di mana ia dapat menyimpan urutan simbol-simbol tumpukan. Kehadiran tumpukan memungkinkan pushdown automaton untuk "mengingat" jumlah informasi yang tak terbatas. Namun, berbeda dengan komputer umum yang juga memiliki kemampuan untuk mengingat informasi sebanyak-banyaknya, pushdown automaton hanya dapat mengakses informasi yang ada di tumpukannya secara cara pertama-masuk-pertama-keluar (last-in-first-out).

Akibatnya, ada bahasa-bahasa yang dapat dikenali oleh beberapa program komputer tetapi tidak dapat dikenali oleh pushdown automaton manapun. Sebenarnya, pushdown automaton mengenali semua dan hanya bahasa-bahasa bebas konteks. Meskipun ada banyak bahasa yang bersifat bebas konteks, termasuk beberapa yang tidak termasuk dalam bahasa-bahasa reguler, ada juga beberapa bahasa yang mudah untuk dijelaskan tetapi bukan merupakan bahasa bebas konteks



Push Down Automata memiliki 7 buah komponen, kita menuliskan spesifikasi PDA P sebagai berikut :

$$P = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$$

- Q adalah himpunan finite state,
- Σ adalah machine alphabet (input symbol),

- Γ adalah himpunan stack symbol,
- δ adalah himpunan transisi,
- q_0 adalah initial state (bagian dari Q),
- Z_0 adalah start symbol untuk initial stack
- F adalah himpunan accepted state

Meskipun PDA secara definisi diperbolehkan untuk bersifat non deterministik, sub kasus yang deterministik sangatlah penting. Terutama, parser (penganalisis) pada umumnya berperilaku seperti PDA deterministik, sehingga kelas bahasa yang dapat diterima oleh automata ini menarik untuk memberikan wawasan terhadap konstruksi mana yang cocok untuk digunakan dalam bahasa pemrograman.

Bahasa HTML (HyperText Markup Language) adalah bahasa yang digunakan untuk membuat halaman web. Bahasa ini terdiri dari tag-tag (elemen) yang memberikan struktur pada konten, seperti teks, gambar, tautan, dan lainnya, sehingga dapat ditampilkan secara terstruktur di dalam browser.

Dengan menggunakan tag-tag HTML, Anda dapat membangun struktur halaman, menambahkan konten seperti teks dan gambar, serta memberikan format dan tata letak pada website.

HTML terdiri dari kombinasi teks dan simbol yang disimpan dalam sebuah file. Untuk membuat file HTML, harus mengikuti standar atau format khusus. Format ini tertuang dalam standar kode internasional atau ASCII (American Standard Code for Information Interchange).

Pada level yang paling dasar, PDA dapat digunakan untuk memeriksa kelurusan struktur tag HTML. PDA akan memiliki state untuk masing-masing elemen HTML, serta aturan transisi berdasarkan tag-tag yang diharapkan dan bagaimana tag tersebut dipasang. Ini memungkinkan PDA untuk mengenali struktur HTML yang valid atau mendeteksi kesalahan sintaksis.

BAB III

IMPLEMENTASI DAN PENGUJIAN

3.1 Spesifikasi Teknis Program

3.1.1 Struktur Data

Struktur Data yang kami gunakan pada aplikasi ini adalah

```
class PDA:
    def __init__(self, states, input_alphabet, stack_alphabet, transition,
start_state, initial_stack_symbol):
        self.states = states
        self.input_alphabet = input_alphabet
        self.stack_alphabet = stack_alphabet
        self.transition = transition
        self.start_state = start_state
        self.current_state = start_state
        self.initial_stack_symbol = initial_stack_symbol
        self.stack = [initial_stack_symbol]
```

3.1.2 Fungsi dan Prosedur

<code>def is_accepted(self)</code>	Fungsi ini digunakan untuk check apakah PDA diterima atau tidak dengan check isi stack terakhir
<code>def process_input(self, input_string)</code>	Fungsi ini digunakan untuk memproses input yang diterima dengan PDA

3.2 Screenshot Hasil Pengujian

3.2.1 Test Case 1

```
<html>
  <head>
    <title>Simple Webpage</title>
  </head>
  <body>
    <h1>Hello, World!</h1>
    <p>This is a simple webpage.</p>
  </body>
</html>
```

Hasil :

Accepted.

3.2.2 Test Case 2

```
<html>
  <head>
    <title>Simple Webpage</title>
  </head>
  <body>
    <h1>Hello, World!<h1>
    <p>This is a simple webpage.</p>
  </body>
</html>
```

Hasil :

Rejected.

3.2.3 Test Case 3

```
<html>
  <body>
    <h1>Hello, World!</h1>
    <p>This is a simple webpage.</p>
  </body>
  <head>
    <title>Simple Webpage</title>
  </head>
</html>
```

Hasil :

TS C. Use
Rejected.

3.2.4 Test Case 4

```
<hmf>
  <head>
    <title>Simple Webpage</title>
  </head>
  <body>
    <h1>Hello, World!</h1>
    <p>This is a simple webpage.</p>
  </body>
</hmf>
```

Hasil :

Rejected.

3.2.5 Test Case 5

```
<html>
  <body>
    <h1>Hello, World!</h1>
    <p>This is a simple webpage.</p>
  </body>
</html>
```

Hasil :

TS C. Use
Rejected.

3.2.6 Test Case 6

```
<html>
  <head>
    <title>Simple Webpage</title>
  </head>
  <body>
    <h1>Hello, World!</h1>
    <h2>Welcome to my page</h2>
    
    <p>This is a <em>simple</em> webpage.</p>

    <div id="footer" class="footer"> This is the end of the page </div>
  </body>
</html>
```

Hasil :

Accepted.

3.2.7 Test Case 7

```
<html>
  <head>
    <title>Simple Webpage</title>
  </head>
  <body>
    <!-- Bagian utama web -->
    <h1>Hello, World!</h1>
    <h2>Welcome to my page</h2>
    <hr>
    
    <p>This is a <em>simple</em> webpage.</p>

    <!-- Custom element -->
    <div id="footer" class="footer"> This is the end of the page </div>
  </body>
</html>
```

Hasil :

Accepted.

3.2.8 Test Case 8

```
<html>
  <head>
    <title>Simple Webpage</title>
  </head>
  <body>
    <!-- Bagian utama web -->
    <h1>Hello, World!</h1>
    <h2>Welcome to my page</h2>
    <img alt="Welcome Banner">
    <p>This is a <em>simple</em> webpage.</p>

    <!-- Custom element -->
    <div id="footer" class="footer"> This is the end of the page </div>
  </body>
</html>
```

Hasil :

PS C:\User
Rejected.

3.2.9 Test Case 9

```
<html>
<head>
  <title>Simple Webpage</title>
</head>
<body>

<h2>HTML Forms</h2>

<form action="/action_page.php" method="POST">
  <h5 class="label">First name:</h5><br>
  <input type="text" id="fname"><br>
  <h5 class="label">Last name:</h5><br>
  <input type="text" id="lname"><br><br>
  <button type="submit">Submit</button>
</form>
```

```
<p>If you click the "Submit" button, the form-data will be sent to a page called "/action_page.php".</p>
```

```
</body>  
</html>
```

Hasil :

Accepted.

3.2.10 Test Case 10

```
<html>  
<head>  
  <title>Simple Webpage</title>  
</head>  
<body>  
  
<h2>HTML Forms</h2>  
  
<form action="/action_page.php" method="TEMLAK">  
  <div id="label">First name:</div><br>  
  <input type="text" id="fname"><br>  
  <div id="label">Last name:</div><br>  
  <input type="text" id="lname"><br><br>  
  <button type="submit">Submit</button>  
</form>  
  
<p>If you click the "Submit" button, the form-data will be sent to a page called "/action_page.php".</p>  
  
</body>
```

Hasil :

Rejected.

3.2.11 Test Case 11

```

<html>
  <head>
    <title>Simple Webpage</title>
    <script>
      document.getElementById("demo").innerHTML = "Hello JavaScript!";
    </script>
  </head>
  <body>
    <h1>The script element</h1>
    <a>Not going anywhere</a><br>
    <a href="https://www.google.co.id/">Might send you somewhere</a>

    <p id="demo"></p>
  </body>
</html>

```

Hasil :

Accepted.

3.2.12 Test Case 12

```

<html>

<head>
  <script>
    document.addEventListener("DOMContentLoaded", function() {
      console.log("Hello");
    });
  </script>
  <link rel="stylesheet" href="path/to/your/styles.css">
</head>

<body>
  <!-- Komentar -->
  <hr>
  <div class="container-flex" style="color: blue; background-attachment:
fixed;" id="divawal">
    <!-- Komentar -->
    <hr>
    <div class="flex-item">
      <hr>
      <p>
        Hello <!-- Komentar -->
      </p>
      <div>
        <h2 class="aaaaa" id="">
          <!-- Komentar -->

```

```

        Pinjam dulu 100
    </h2>
</div>
</div>
<hr>
<a href="#"></a>
</div>
<form id="adsad" method="POST">
    <h1>asdadasd</h1>
    <p style="adasadad;"></p>
    <br>
    <div class="asdsda">
        <h3>
            dasdaasdda
        </h3>
        <div style="asdaasdadas;"></div>
    </div>
</form>
</body>

</html>

```

Hasil :

Accepted.

3.2.13 Test Case 13

```

<html>
<head>
    <script>
        window.onload = function() {
            alert("Hello from nested script");
        };
    </script>
    <link rel="stylesheet" href="styles.css">
</head>
<body>
    <!-- Outer container -->
    <div class="outer-container">
        <h1>Welcome to the Nested World</h1>

        <!-- Nested container with a form -->
        <div class="nested-container">
            <form id="nested-form" method="POST">

```

```

        <input type="text" id="username">

        <input type="password" id="password">

        <button type="submit">Submit</button>
    </form>
</div>

<!-- Another nested container with a list -->
<div class="another-nested-container">
    <h2>Items List</h2>

</div>

</div>
</body>
</html>

```

Hasil :

Accepted.

3.2.14 Test Case 14

```

<html>
<head>
    <script>
        window.onload = function() {
            alert("Hello from nested script");
        };
    </script>
    <link rel="stylesheet" href="styles.css">
</head>
<body>
    <div>
        <div>
            <div class="div1">
                <form>
                    <h4> Ini diisi dalam h4
                    <p>
                        <small>
                            asdasddsda
                        </small>
                        <em>

                        </em>
                    </p>
                </h4>
            </form>
        </div>
    </div>

```

```

    </div>
  </div>
  <script>
    window.onload = function() {
      alert("Hello Ini script kumpulan div ngeri");
    };
  </script>
</div> kelebihan /div

</body>
</html>

```

Hasil :

ERROR: NO
Rejected.

3.2.15 Test Case 15

```

<html>
<head>
  <script>
    window.onload = function() {
      alert("Hello from nested script");
    };
  </script>
  <link rel="stylesheet" href="styles.css">
</head>
<body>
  <div>
    <div>
      <div class="div1">
        <form>
          <h4> Ini diisi dalam h4
            <p>
              <small>
                asdasdddsda
              </small>
              <em>

            </em>
            </p>
          </h4>
        </form>
      </div>
    </div>
  </div>
  <script>

```

```
        window.onload = function() {  
            alert("Hello Ini script kumpulan div ngeri");  
        };  
    </script>  
</body>  
</html>
```

Hasil :

Accepted.

BAB IV
PEMBAGIAN TUGAS

NIM	NAMA	PEMBAGIAN TUGAS
13522129	Hugo Sabam A	pda,test case,diagram,laporan,pytho n
13522136	Muhammad Zaki	python,diagram,laporan,pd a
13522138	Andi M Sitorus	pda,test case,diagram,python

LAMPIRAN

Link repository :

<https://github.com/miannetopokki/HTMLCheckerPDA>

Link Diagram State :

<https://drive.google.com/file/d/1Sc7-mtOFoSjm1dxI1DZzXpUhbMydhi5w/view?usp=sharing>