

SI 413 Fall 2021: Homework 6

Due Friday, October 6

Your name:

Citations and collaborators:

Comments, suggestions, or questions for your instructor:

Fill out the first row of the table on a 0-5 scale before turning in.

This rubric is also available on the website under “Admin”:

- **5:** Solution is completely correct, concisely presented, and neatly written.
- **4:** The solution is mostly correct, but one or two minor details were missed, or the presentation could be better.
- **3:** The main idea is correct, but there are some significant mistakes. The presentation is somewhat sloppy or confused.
- **2:** A complete effort was made, but the result is mostly incorrect.
- **1:** The beginning of an attempt was made, but the work is clearly incomplete.
- **0:** Not submitted.

Problem	1	2	3	Total
Self-assessment				
Final assessment				

1 Fixing a grammar

The following grammar defines a language for (opening) HTML tags, like `<input checked type="checkbox">`

```
start → LA inner RA
inner → NAME attrs
attrs → attrs single
attrs → ε
single → NAME EQ VAL
single → NAME
```

Fix this grammar (without modifying the language it accepts!) to make it LL(1). Write your new, updated grammar below.

2 Top-down parsing

Using your LL(1) grammar from the previous question, show the *partial parse tree* that would result from reading the following sequence of tokens, in a top-down parser.

LA NAME NAME NAME EQ VAL

Note, this sequence does not parse all the way up to the start symbol. So it will be a *partial* parse tree, with some items un-expanded or un-matched.

3 Bottom-up parsing

Now use the *original* grammar from problem 1 to complete a partial parse of the same sequence of tokens, but this time using a bottom-up parsing strategy. Here are the grammar and tokens again:

```
start → LA inner RA
inner → NAME attrs
attrs → attrs single
attrs → ε
single → NAME EQ VAL
single → NAME
```

```
LA NAME NAME NAME EQ VAL
```

Again, this will be a *partial* parse. But for bottom-up parsing, that means we will have a “forest” of multiple trees, that haven’t yet combined all the way to get back to the start symbol.

Assume the next token of look-ahead is **RA**, in order to combine the partial parse trees (reduce) as much as possible.