

**Sabancı University**  
**Faculty of Engineering and Natural Sciences**

**CS305 Programming Languages**

**Homework 2**

Due: March 11, 2019 - Monday @ 23:55 (upload PDF to SUCourse)

## 1 Introduction

In this homework you will write a context free grammar for CSML language for which you designed a scanner in the previous homework. Note that, there might be differences between the syntax of CSML language given in the previous homework and this one. Therefore, take the explanations on the syntax of CSML language given in this homework.

You will write a CFG and extended BNF notation is not allowed. The language that will be generated by your grammar and other requirements of the homework are explained below.

## 2 The Language

The grammar you will write will generate the CSML language as described below. Here is an example program in this language to give you an idea how a CSML program looks like.

```
<course code="CS305" name='Programming Languages' type="Lecture">
  <class section="0" instructor='Husnu Yenigun' crn=20258 capacity=60>
    <meeting day=R start=08:40 end=10:30/>
    <meeting start=08:40 end=09:30 day=F/>
  </class>
</course>
<course code="CS301R" name="Algorithms-Recitation" type="Recitation">
  <class section="0" instructor='Husnu Yenigun' crn=20257>
    <meeting start=17:40 day=M end=18:30/>
  </class>
</course>
<constraint>
  <item code="CS305"/>
  <item crn=20257/>
</constraint>
```

Below is the detailed syntactic features of the CSML language.

1. A CSML *program* consists of a list of top level *elements*.
2. The sequence of top level elements maybe empty, or contains one or more top level elements.
3. A top level element is either a *course element*, or a *constraint element*.
4. A *course element* consists of a *course opening tag* and a *course closing tag*. Between a *course opening tag* and a *course closing tag*, there is a non-empty list of *class elements*.
5. A *course opening tag* consists of a `tOPEN` token, `tCOURSE` token, a non-empty list of *course attributes* and a `tCLOSE` token.
6. A *course attribute* gives either
  - the code of the course (e.g. `code="CS305"`), or
  - the name of the course (e.g. `name="Programming Languages"`), or
  - the type of the course (e.g. `type="Lecture"`)

Note that, your grammar should allow these attributes to be given in any order, it should not force every attribute to be used (i.e. it is okay if some attributes are missing), and it should not restrict the multiple use of an attribute (i.e. it is okay if an attribute is seen more than once).

7. A *course closing tag* is simply `</course>`.
8. A *class element* consists of a *class opening tag* and a *class closing tag*. Between a *class opening tag* and a *class closing tag*, there is a non-empty list of *meeting elements*.
9. A *class opening tag* consists of a `tOPEN` token, `tCLASS` token, a non-empty list of *class attributes* and a `tCLOSE` token.
10. A *class attribute* gives either
  - the section information for the class (e.g. `section="0"`), or
  - the instructor information for the class (e.g. `instructor="Husnu Yenigun"`), or
  - the CRN of the class (e.g. `crn=20258`), or
  - the capacity of the class (e.g. `capacity=60`)

Note that, your grammar should allow these attributes to be given in any order, it should not force every attribute to be used (i.e. it is okay if some attributes are missing), and it should not restrict the multiple use of an attribute (i.e. it is okay if an attribute is seen more than once).

11. A *class closing tag* is simply `</class>`.
12. A *meeting element* consists of a `tOPEN` token, `tMEETING` token, a non-empty list of *meeting attributes* and a `tSELF` token.
13. A *meeting attribute* gives either
  - the start time information for the meeting (e.g. `start=17:40`), or
  - the end time information for the meeting (e.g. `start=18:30`), or
  - the day information for the meeting (e.g. `day=M`)

Note that, your grammar should allow these attributes to be given in any order, it should not force every attribute to be used (i.e. it is okay if some attributes are missing), and it should not restrict the multiple use of an attribute (i.e. it is okay if an attribute is seen more than once).

14. A *constraint element* consists of *constraint opening tag* (which is simply `<constraint>`) and a *constraint closing tag* (which is simply `</constraint>`). Between a *constraint opening tag* and a *constraint closing tag*, there is a non-empty list of *item elements*.
15. An *item element* consists of a `tOPEN` token, `tITEM` token, an *item attribute* and a `tSELF` token.
16. An *item attribute* gives either
  - a code (e.g. `code="CS305"`), or
  - a CRN (e.g. `crn=20257`)

### 3 Terminal Symbols

Use the following terminal symbols in your grammar. You can assume that the scanner will return the corresponding token. Do not add any new tokens and do not change the name of the tokens.

Token	Lexeme	Token	Lexeme
tOPEN	<	tCLOSE	>
tEND	</	tSELF	/>
tCOURSE	course	tCLASS	class
tMEETING	meeting	tCONSTRAINT	constraint
tITEM	item	tNAME	name=
tCODE	code=	tINSTRUCTOR	instructor=
tTYPE	type=	tCRN	crn=
tSECTION	section=	tCAPACITY	capacity=
tSTART	start=	tEND_A	end=
tDAY	day=	tSTRING	anything in quotations
tMON	M	tNUM	any positive integer
tTUE	T	tWED	W
tTHU	R	tFRI	F
tTIME	Any time between 00:00 and 23:59		

Besides these tokens, you if you need any more tokens in your grammar (e.g. [, ;, etc.), directly use the lexeme of such tokens in your grammar.

## 4 Non-Terminal Symbols

Use the following non-terminal symbols in your grammar. This is the entire set of non-terminals that you will use in your grammar.

- Do not add a new non-terminal symbol
- Do not change the name of a non-terminal symbol
- Do not remove any non-terminal symbol

**prog**: This is the start symbol of the grammar.

**elementList**: Denotes a list of elements.

**element**: Denotes an element.

**beginCourse**: Denotes a begin tag for courses.

**endCourse**: Denotes an end tag for courses.

**beginConstraint**: Denotes a begin tag for constraints.

**endConstraint**: Denotes an end tag for constraints.

**courseAttrList**: Denotes a list of course attributes.

**courseAttr**: Denotes a course attribute.

**classList**: Denotes a list of classes.

**class**: Denotes a class.

**beginClass**: Denotes a begin tag for classes.

**endClass**: Denotes an end tag for classes.

**classAttrList**: Denotes a list of class attributes.

*classAttr*: Denotes a class attribute.  
*meetingList*: Denotes a list of meetings.  
*meeting*: Denotes a meeting.  
*beginMeeting*: Denotes a begin tag for meetings.  
*endMeeting*: Denotes an end tag for meetings.  
*meetingAttrList*: Denotes a list of meeting attributes.  
*meetingAttr*: Denotes a meeting attribute.  
*day*: Denotes a day tag.  
*itemList*: Denotes a list of items.  
*item*: Denotes an item.  
*beginItem*: Denotes a begin tag for items.  
*endItem*: Denotes an end tag for items.  
*itemAttr*: Denotes an item attribute.

## 5 How to Submit

On SUCourse, submit an PDF document that contains the grammar you have written. Name the document as id-hw2.pdf where id is your student ID.

## 6 Notes

- **Important:** SUCourse's clock may be off a couple of minutes. Take this into account to decide when to submit.
- No homework will be accepted if it is not submitted using SUCourse.
- You must write your files by yourself.
- Start working on the homework immediately.