

Tutorial for University Entrance Exam Task

事前準備

まずシステムをつくる上で必要なファイルをダウンロードする必要がある。

アカウントを申請して、許可がおりたあとにログインして、必要なファイルをマイポータルのページの下に「問題・正答 XML ファイルダウンロード」のリンクからダウンロードする。

重要ファイルの概要

ダウンロードした問題・正答 XML の zip ファイルを展開すると、“Center-XML” のフォルダ以下にいくつかのファイル・フォルダができる。そのうち特に重要なものは以下の通りになる。

Center-XML-dtd

センター試験の問題 XML ファイルの入力形式が定められている。試験問題に対するアノテーションの具体的な仕様は Doc の下のファイルに詳細が書いてある。

Center-XML-dev

センター試験の訓練データが教科ごとに分かれて入っている。システムをつくる際に必要なファイルだけ参照すればいい。

Seitouhyou

訓練データの解答 XML ファイルとその形式ファイルがおいてある。

Center-XML-test

テストデータの問題ファイルがおいてある。

システムの作り方

問題構造 XML の解析

詳細は仕様ファイルに書いてあるが、必要最小限のものについて、1993 英語本試験を例に説明する。

XML ファイルに exam タグの下に、大問 question が 6 個あり、それぞれの大問の中に小問 question があり、さらにその下に問が入っているものもある。入力を受け取るときに question の一番下の階層まで解析して、以下のようなものになったとする。

```
<?xml version="1.0">
...
<exam source="..." ...>
...
  <question id="Q1" minimal="no">
    ...
    <question anscol="A50" answer_style... answer_type=... id=... knowledge_type=... minimal=...>
      <choice ansnum="1">...</choice>
      <choice ansnum="2">...</choice>
      <choice ansnum="3">...</choice>
      <choice ansnum="4">...</choice>
    </question>
  </question>
</exam>
```



ここで **anscol** はそれぞれの解答欄に割り当てられるユニークな番号で、解答時にこの番号と答えを両方出力する必要がある。さらに、それぞれの選択肢に対応するものは **ansnum** である。回答するときこの値を出力する。

問題を解く

それぞれの小問に対して、それぞれの工夫で正しい choice を選択すればよい。たとえば上記の問題に対して、ランダムで出力するとしたら、c 言語だと以下のようにかける。

```
int solution_func(Question *q){
    int system_answer = 1+rand()%4;
    return system_answer;
}
```

出力の仕方

解答を提出するとき、正答表と同じ XML 形式で提出する必要があるが、正答表のすべてのフィールドが必要になるわけではない。

解答全体を<answerTable>ルート要素の下におき、それぞれの最小の問に対し<data>で解答する。<answerTable>タグの file 属性には対象の試験問題のファイル名(拡張子を除く)を指定する。回答するとき、最小限のものは<answer>と<anscolumn_ID>である。<answer>は上の ansnum に対応し、<anscolumn_ID>は上の anscol に対応する。さらに answerTable の属性指定と XML のヘッダーの二行を入れる必要があり、この二行に関しては以下の出力例のようにすればよい。この例では、上の問いに対して選択肢 1 を選択している。

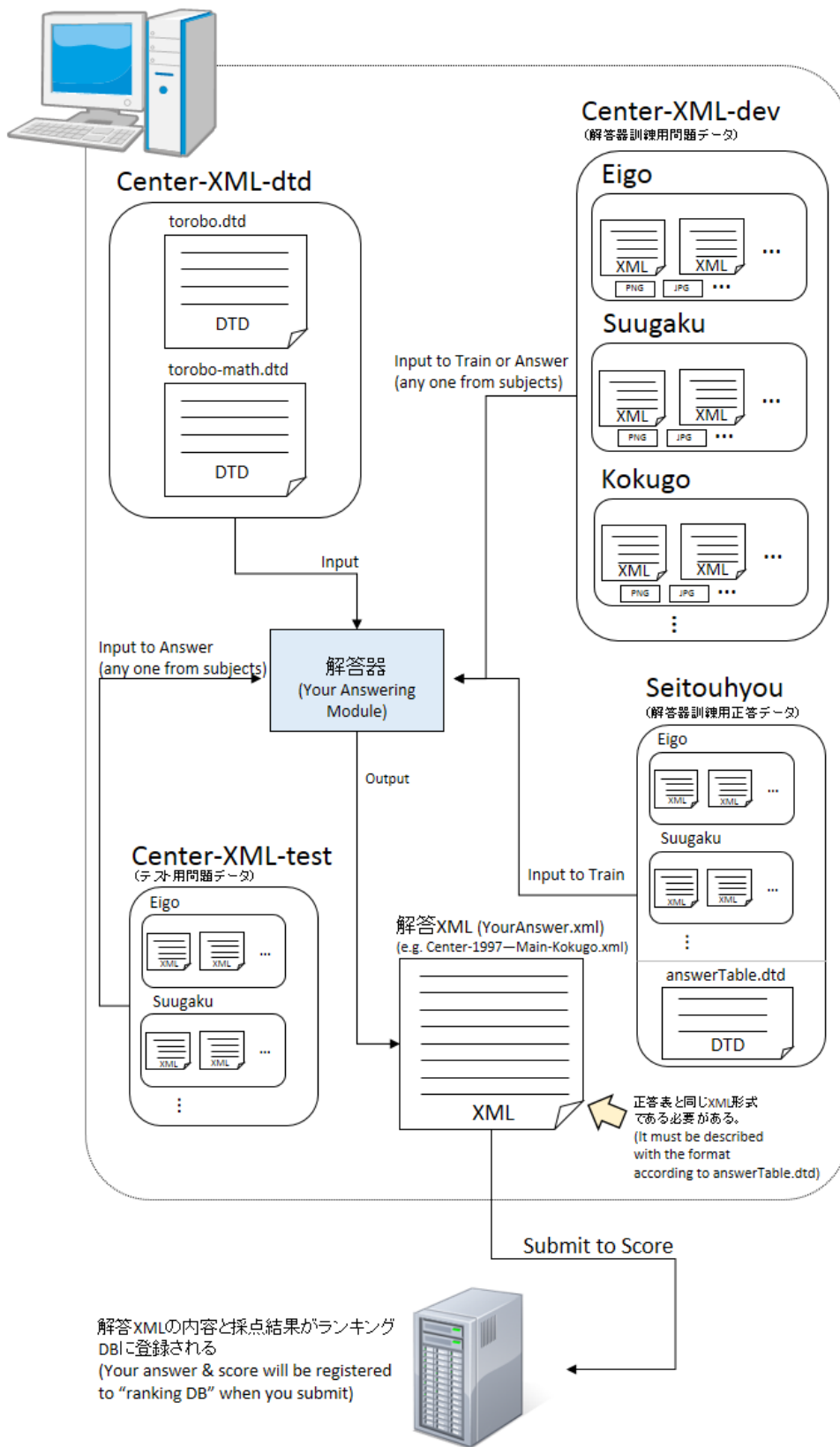
```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE answerTable SYSTEM "http://21robot.org/answerTable.dtd">
<answerTable file="Center-1993--Main-Eigo">
    <data>
        <answer>1</answer>
        <anscolumn_ID>A50</anscolumn_ID>
    </data>
</answerTable>
```

問題を解けない場合はそれに対応する<data>を作らなければよい。

もっと詳しい出力の仕方は Doc にあるファイルを参照。

解答提出

解答の XML ファイルの名前は正答表と同じ形式でつける。たとえば、1997 年国語本試験に対して、Center-1997--Main-Kokugol.xml とすればよい。提出するときはマイポータルの“解答アップロード”を使ってアップロードすると、システム判定がなされる。



NII 東ロボ事務局 採点サーバ
(Scoring Server@Todai Robot Project Office, NII)

図：解答器の訓練、解答データの生成、解答データの採点・登録の流れ