

FIRST - Bug #4058

Question 46 on Fall05-Exam3.pdf causes error in recognition

05/11/2009 07:35 AM - Ryan McFall

Status:	New	Start date:	05/11/2009
Priority:	Normal	Due date:	
Assignee:	Sandeep Namilikonda	% Done:	0%
Category:	parser	Estimated time:	0.00 hour
Target version:	client-prototype	Spent time:	0.00 hour
Bugzilla-Id:	4058		
<p>Description</p> <p>(This is a resubmission of a bug lost due to fire. The old bug number was 4057; that number seems to have taken by a newer bug submitted after the restore. We talked about this bug during the videoconference on 2009-05-06).</p> <p>Question 46 on the exam private/samples/Fall05-Exam3.pdf is not recognized correctly. This question contains some text, followed by a line of text on it's own line that represents a string of DNA, and then contains several choices which are all numbers.</p> <p>Not all of the prompt is put into the question text, and the last choice (E) also contains the question text for the next question.</p> <p>Sandeep, can you investigate this and make a report back as to why this is occurring?</p>			

History

#1 - 05/13/2009 09:19 PM - Sandeep Namilikonda

The following pseudo code is relevant in the context of Bugs 3923, 3924, 3989, 4058, 4005. I tried to elaborate more for the conditions that seemed more probable to be faulty.

```
PDFAssessment: parse() {  
  
  // SeparateQuestions()  
  
  // AssociateImages()  
  
  // CombineSplitQuestions()  
  
  // RecognizeQuestions()  
  
}  
  
Pseudo code for SeparateQuestions():  
  
for each string block "i":  
  
  1) Find new section or part  
  
  2) In the text block, if found  
  
    (i) delimited question  
  
    (ii) "answer key" text  
  
    (iii)instr block and range  
  
    (iv) image tag  
  
    (v) digit  
  
  - if expected question # found => set nextNumMatches and  
    update qnum list and the index (multi-digit case)!  
    - else if new section found  
    - else if answer key found {} (goto for "i")!
```

```
- else append text to current question {} (goto for "i")!
```

```
- if correct number found then
```

```
(below are examples of cases handled where  
    134 is an arbitrary question number)
```

```
exclude 134" " and 134"a"  
    exclude 134."a" and 134.")" and 134."$"  
    exclude 134" " and 134" a"
```

```
allow:
```

```
    134)  
    134" ".  
    134" "$  
    134#  
    134 a  
    134 #
```

```
- check for answer key "<number>;) <letter>";  
    - if index+1 is digit then append to current Q  
    - else set current Q's question number and prepend  
      instr block if any!
```

```
- else if new page then change pagenum and set question to current  
    - else if left or right indented text  
      - if right then append
```

```
(vi) instr block
```

```
- update instr block with current  
    - Remove the text from list and i--.
```

```
(vii)white lines between two consecutive lines of text
```

```
- if text corresponds to instructions (e.g., questions i-j)  
    then choose the text to be the new question text and set  
    indented to false.  
    - else append current text to prev question and add new box  
      to existing one. Remove the text and i--.
```

```
(viii)else same question but check indentation
```

```
- if left indented and choice text found (e.g., a. or b) )  
    then add the current text to prev question and update box.  
    - else set indented to true if current.x > q.x.  
      Add the current text to prev question and update box.  
    - Remove text and i--.
```

Most of the cases are handled by this last condition! (e.g., choices, multiple lines of question). Infact, this is what results in bug 4058 due to faulty indentation check!

#2 - 03/27/2013 02:25 PM - Redmine Admin

Original Bugzilla ID was 4058