AUTOMATIC GENERATION
OF MULTIPLE-CHOICE TESTS:
THE PAST, THE PRESENT AND THE FUTURE

Ruslan Mitkov

Abstract. *Multiple-choice tests are a popular form of objective assessment in which respondents are asked to select only one answer from a list of choices, and are extensively used in teaching, learning, training, market research, elections and TV shows. This keynote speech will start with a brief review of the history of multiple-choice tests, illustrating their diverse use.*

The presentation will then provide details as to how automatic generation of multiple-choice tests started, more specifically going back to the late 1990s when the speaker for the first time proposed Natural Language Processing (NLP) methodology to automate (and enhance the efficiency of) the task of multiple-choice tests construction, a labour-intensive and time-consuming process. The speaker will then outline the pioneering work on this topic by Mitkov and Ha (2003) which he presented at a conference in Edmonton, Canada, and share the story how the presentation of this novel technology led to a multi-million-dollar project funded continuously since 2005 by a major US organisation.

Next, the presentation will go into a slightly more technical phase, outlining the core of Mitkov's NLP methodology with the selection of distractors being an especially interesting semantic similarity task and illustrating this methodology through several easy-to-follow examples. While NLP cannot perform magic and is prone to errors, the presenter will explain that even with a number of automatically generated texts having to be discarded and the majority having to be post-edited, the semi-automatic approach (the machine generating a very high number of possible multiple-choice tests and the human post-editing), does still lead to a much improved time efficiency without the quality being compromised. The speaker will back up these claims with the results of evaluation experiments.

Mitkov and Ha's (2003) pioneering work led to a large body of work on the generation of multiple-choice items which the speaker will briefly review. Recent work includes the employment of Deep Learning techniques.

The speaker will proceed to outline his latest research on the topic which benefits from the latest Deep Learning models to generate questions from more than one sentence. The generation of multiple-choice questions from more than one sentence is a significant breakthrough given that all previous related work had attempted generation from one sentence only.

In the last part of his presentation, the speaker will put forward his views as to what would make exciting next steps in automatic generation of multiple-choice tests.