

Review of the dissertation thesis

Author: **Mgr. Michal Bída**

Title: **Artificial Emotions in Virtual Storytelling**

Supervisor: **Mgr. Cyril Brom, Ph.D.**

Michal has addressed a tough problem in his thesis: that of modeling **semi-interactive narratives** in 3D virtual environments that feature **3D human-like agents expressing emotions**. Ever since Činčera's Kinoautomat, a handful of researchers and artists addressed this issue and similar ones; with minor progress and without substantial breakthroughs.

The key subproblems addressed in the thesis are:

- a) developing a new **control architecture for 3D agents**; an architecture that would be “**minimalistic**” (i.e., without features unnecessary for end-products developers implementing mid-size interactive dramas), yet connect classical agent decision making, steering, and emotion modeling components;
- b) developing a new **method for semi-automatic estimating quality of narrative** enacted by 3D characters in a 3D world.

After an enthusiastic introduction, Michal describes his EWA architecture in Chapter 2 (addressing Point (a)). Thereafter, he evaluates it on two proof-of-concept mini-games called SimDate Level 1 and SimDate Level 2 (Chapters 3.3, 3.4). Though the architecture is well thought off, it is not really a substantial innovation to existing approaches. The author is aware of this and rightly remarks that the key point of the architecture is actually to address Subproblem (b). This happens in **Chapter 4, wherein the actual contribution of the thesis lies**.

When designing a method for semi-automatic estimating quality of narratives, one needs domains on which to test his or her approaches. Based on the domains, one has to generate hundreds of narrative examples and classify them on the high quality–low quality axis. These domains are very rare. Therefore, Michal has used his EWA-based mini-games to serve as these domains/generators of narratives. He further complemented his SimDate-based domains with the MOSS domain by other authors.

The ideas behind Michal's story-quality-estimating approach (to abstract the evolution of the state of a story as a tension between characters, and to detect subscenes in the overall narrative) are not new, but to my knowledge, it is new to use this “tension curve” and the list of subscenes to attempt at clustering stories to “good” and “bad” ones. Michal described his method, implemented it and performed a set of experiments with several metrics. Despite no breakthrough findings (these were actually not to be expected given quite limited previous research in this area), experimental results are encouraging and this part of Michal's work can thus be viewed as first steps in mapping terra incognita of the problem of semi-automatic estimating qualities of narratives.

I'd like to ask Michal: If someone wants to continue in this endeavor, what would be the single most important next step?

Generally, the amount of work done by Michal in this thesis is adequate. The literature review is solid and up-to-date, but few bits seem to be missing.

For example, I'm wondering, to what extent Tactical Iraqi's architecture addresses the problems mentioned in the thesis (R1 – R7, page 38)?

Also, the author's own work (Points (a) and (b) above) is not always contextualized within the literature review (i.e., why and how is Michal's work different from related work is not always stated).

As concerns the text, it is vivid and readable. It would benefit from proof-reading here and there, but somewhat more problematic is that the level of description detail sometimes varies between sections (e.g., cf. 1.5.4 with 1.5.2 and 3.3 with 3.4) and, at some points, the text leaves the reader starving for more details (e.g., Sec. 3.1 [Pogamut's description], 3.4 [SimDate Level 2 description]). Also, reshuffling portions of the text would further improve readability (e.g., the MOSS system is introduced in Sec. 4.5, but its representation is discussed in 4.4).

Related to that, I'm wondering:

-- how exactly SimDate Level 1 and SimDate Level 2 evaluate whether ALMA and Steerings are appropriately integrated within EWA architecture (Sec. 3.3, 3.4)?

-- how are hierarchical goals put forward by EWA evaluated (if at all)?

-- the underlying emotion model ALMA works with positive and negative emotions, which are reduced to a uni-dimensional feeling variable in EWA (page 58, Eq. 2.1); how is it accomplished that the contributions from positive versus negative emotions are balanced, i.e., the feeling variable does not converge to either the minimal or maximal value?

-- what type of stories can be found in "quality" clusters generated by the story-quality-estimating method?

All in all, this is a decent thesis with some non-fatal issues. Especially the methodology for semi-automatic estimating the quality of virtual drama, albeit just a provisional step in addressing a large and considerable problem, is a notable contribution to the interactive digital storytelling field.

I **recommend** the work for defense and I believe the author **should be awarded** the Ph.D. degree.

Mgr. Cyril Brom, Ph.D.

Department of Software and Computer Science Education

Faculty of Mathematics and Physics

Charles University in Prague