GALA 2007 submission document

Title:	DEIRA, Dynamic Entertaining Intelligent Reporter-Agent
Track:	Student
Category:	Race Reporter
Authors:	**The Prancing Ponies** (*team name*)

Almer S. Tigelaar (*contact*)
2nd year Master Student
HMI, CS, Twente University, The Netherlands

Thijs Alofs
2nd year Master Student
HMI, CS, Twente University, The Netherlands

Francois Knoppel
2nd year Master Student
HMI, CS, Twente University, The Netherlands

Danny Oude Bos.
2nd year Master Student
HMI, CS, Twente University, The Netherlands

Contact data:
    name: Almer S. Tigelaar
    e-mail:
    ~~**phone:**~~ ——————————
    postal address:

URL: http://www.queequee.net/deira/

Movie file submitted: deira.avi

Reference teacher:
    Dr. Z. M. Ruttkay
    affiliation: HMI, CS, Twente University
    e-mail: zsofi@ewi.utwente.nl

Process of development: The first prototype of DEIRA was developed for a university master course called *HMI Project* which ran from the beginning of September until the beginning of February. After this project, the team decided to continue development for a second improved prototype to send in for the GALA Challenge. All team members were involved during the entire period.
The goal for the first prototype was to create a working agent, based on preferences from horse race fans obtained prior to development. For the second prototype the main improvements were in the expression of emotions and in the size of the reporter's vocabulary.

Resources used:
1. Visage Interactive, for agent visualization
2. Realspeak, as text-to-speech generator
3. Blender, for face generation and adjustment, using a FaceGen model.
4. Java (editor: Eclipse), for programming

Resources required: DEIRA was developed and tested on a PC with the following specs: AMD Athlon XP 2600+, 1024MB RAM, ATI Radeon 9800 (128MB), running Windows 2000. Minimum requirements not tested.

GALA 2007 submission document

DEIRA, *Dynamic Entertaining Intelligent Reporter-Agent*

1. The application and context of the work

Application context   DEIRA is a race reporter, specifically designed for reporting on horse races. The intention is to provide an alternative for real race reporters.

Usage and adaptation   The system may be adapted in many ways. The face can easily be replaced, as long as it can be used by Visage Interactive. The voice can be generated by any text-to-speech generator. What events are generated and the corresponding sentences that can be uttered about that event can also easily be adjusted, since they are expressed in rules and grammars respectively. This means adapting the reporter to report on a different type of race should be an easy thing to do. Also, the system is designed in a very modular way, which should make other structural changes fairly straight-forward to make.

User testing   For the initial prototype, user opinions were obtained multiple times during the development process. The second prototype is based mainly on the final user feedback on the first prototype. The second prototype has not been tested with users explicitly.

2. Novelty

Our agent has a large vocabulary, which is easily extendible to include even more. Emotional display is done by intonation of speech as well as physically in the face. The design was highly influenced by the opinions of potential users (in the case of horse race reporting), who were polled multiple times during development. The system is also easily adjustable for use in other kinds of races.

3. The architecture

General architecture   The provided horse race simulation software is used to provide input on the state of the race. DEIRA analyses this information and eventually sends the resulting utterance including emotional expression to the visual representation software Visage Interactive which in turn shows the agent reporting the race.

Within DEIRA, horse race information is received by the *Race Analysis Module*, which scans the changing race state for events that are possibly interesting to report. These events are then passed on to the *Mental Model* which adjusts the global mental state and adds a local mental state to the incoming events. The events are then queued in the *Event Queue*, making them available for the other modules. The *Text Generation Module* takes the highest priority event without text and generates suitable sentences based on the grammar to describe it. The *Speech Generation Module* is responsible for making this text suitable for the text-to-speech engine, also adding the emotional aspects to the speech. The *Face Animation Module* provides the corresponding facial expression information. When an event has obtained all the necessary information from these modules, the *Output Module* sends it all through to Visage.

Modalities of output   Verbal (speech), nonverbal (display of emotions on the face)

Report management   Events are assigned an importance value, which decays over time depending on the type of event. The most important event will be reported, and as long as there are still events in the queue, there is something for the reporter to say. As of yet, there are no silence fillers.

Additional aspects   A global emotional state of the reporter is maintained (which changes with the events occurring in the race), as well as a local emotional state relating to specific events. The basic personality of the reporter is stored as parameters in the personality database.

4. Performance

Pros

- Large vocabulary: based on the utterances generated for just one race, it can be said that, while not taking into account the variety caused by the different horse names, at least 1030 different sentences can be spoken by the reporter. Apart from certain events that have not happened in the specific race used to obtain this statistic, this number also excludes variety in speed, pitch and voice.
- Expressive face
- Easily extendible and adjustable in many aspects
- Based on user preferences

Cons

- Artificial speech
- Limited head movement
- Does not use background information or fillers during silent periods

Reporting delay

The time it takes for an event from creation to passing it on to Visage is 25ms at the most. The sending/receiving delay is also 25ms in the worst case. This means this system does not add more than 50ms to the total delay from the occurrence of an event to the reporting of it. There is also additional delay on the horse race simulation software side, on which we have no influence.

User responses

User evaluation for the first prototype indicated that while users appreciate the basic functionality, the expression of emotion and the variety in speech are very important aspects as well. For the second prototype these were the main targets of improvement. The system as it is now has not been evaluated with users as of yet.

The Prancing Ponies *(left to right: Thijs, François, Danny, Almer)*