Multimedia Aware Discourse Generation
Justus Robertson, Colin Potts, R. Michael Young
Liquid Narrative Group

Problem
- Generate multimedia summarizations of game play traces from DotA 2, a popular MOBA game.
- Intelligently communicate information using a combination of dynamically generated text, maps, and cinematics. Minimize redundant information by reasoning across independent mediums.

Generating Text
Text generation takes place in three phases:
- A discourse plan is created. The planner introduces characters and objects, chooses actions and states to communicate, and chooses orderings between utterances based on communicative goals.
- The discourse plan is translated to an RST-like tree. From the plan a tree structure is generated that represents relations between different utterances.
- A micro-planner takes the tree as input and produces text.

Generating Film
A film is a series of camera shots of subjects in the game world over a span of time. Planner has a library of types of camera shots to choose from when creating a film.

Shot library includes:
- Shot types: Close up, medium shot, full shot, long shot, extra long shot, over the shoulder shot, and two shot.
- Camera action types: dolly, pan, crane, tilt, focus.
- Camera angle types: low, medium, high.
- Different camera focal lengths and apertures.

Camera shots of particular subjects and particular times are compiled in film objects to convey communicative goals. Once the shot sequences are planned, they are converted to an intermediary XML representation, called Oshmirto, and sent to the film realizer, called FireBolt, for rendering.

Generating Maps
Maps are generated to quickly show an overview of events that have taken place in the DotA 2 play trace. The map images contain a legend that describes map icons, a large map image that may show a small portion of the overall game world, and a smaller summary context map that shows the entire context of the game world.

The discourse planner can add three types of elements to a map object:
- Character positions
- Different varieties of attacks
- Character movements from one place on the map to another

Multimedia
Multimedia discourse plans are created using both abstract and atomic text, map, and film actions.
- The planner uses two action types:
  - High level abstract actions. These quickly add common utterance sequences to the discourse plan.
  - Low level atomic actions. Utterance sequences ground out into atomic actions that create new text, map, or film objects or add specific communicative events to a particular object.
- The result of multimedia planning is a plan that contains integrated text, maps, and film.

Defense of the Ancients 2
Defense of the Ancients (DotA) 2 is a multiplayer online battle arena (MOBA) game we use as an example domain for the summarization system.

DotA 2 information:
- Online multiplayer 5 vs. 5 game
- Games last 45 to 60 minutes
- Players compete to gather resources and destroy opposing players and bases
- We convert DotA gameplay traces to XML documents
- XML consists of logical sentences about the game world