

NOTES FROM THE DISCUSSION GROUP ON “EVALUATION”

Members: Harold Cohen <hcohen@ucsd.edu>
Maggie Boden <M.A.Boden@sussex.ac.uk>
Dave Brown <dcb@cs.wpi.edu>
Paul Brown <paul@paul-brown.com>
Oliver Deussen <deussen@inf.uni-konstanz.de>
Philip Galanter <galanter@viz.tamu.edu>

Note: These notes represent approximately what was discussed by the group members over a period of several hours over two days. There has been some attempt to organize the material, but little attempt to expand it to make it coherent—we rambled, so do the notes. The notes, and this report, were recorded, organized, and elaborated into this form by Dave Brown.

Version: Fri Aug 7 20:26:08 EDT 2009

FOCUS: Evaluation of artistic artifacts

FRAMEWORK:

Point of view:

- as viewer/experiencer;
- as creator;
- as interactive participant.

What aspects of evaluation can be made computational?

Does machine need to mimic the evaluation methods of humans?

Evaluation of partially complete artifacts?

Evaluation of complete artifacts?

- When is a work of art finished?

Evaluate artifact or process used to produce it?

Evaluate for “quality”

Evaluate for “value”

- not just \$ value
- depends on the “consumer”

Evaluation can feed back into the system to affect (hopefully improve) future performance.

Harold's desire is to have a system that decides, as he does, which of AARON's generated images are worth saving and perhaps presenting to the public.

If evaluation during generation is in the same terms (i.e., uses same "rules") then there is a danger of not being transformationally creative.

DECISION IMPACT:

What is the impact of initial decisions during an artifact's creation on the final quality?

Are these decisions reversible? (e.g., pencil drawing vs. carving)
What effect does that have final quality?

For products initial decisions can make major effects.
Some artistic artifacts are similar to products in that regards, others not.

Cezanne is reported to throw paintings away once an "incorrect" brushstroke was made.
i.e., reversibility impacted by evaluation.

LEVELS OF EVALUATION:

Evaluation at the hand-eye level -- e.g., "quality of line".

Different points of view for evaluation

- e.g., quality of "marks";
 - innovative use of media;
 - technical quality;
 - ability to tell a story.

Evaluation of "consumer's" emotional reaction.

- from facial expressions;
- from expected reaction to content (e.g., to Munch's "The Scream")

Novelty and surprise are included in most evaluations of the creativity of an artifact.

Can also evaluate for artistic "style" (e.g., cubist)

Visual arts: formal evaluation, based on principles.

- closure (closed shapes);
- foreground/background;
- balance;
- inside/outside;
- bottom up evaluation.

Can principles be obtainable by machine learning?

NOVELTY:

Novelty: of different attributes of the artifact

- e.g., color, density, texture, material, etc.
- of concepts, intentions, ...
- hierarchical evaluation possible
- can accumulate
- which attributes are significant?

Novelty in art: of the media and/or the “meaning”

1. Novelty *detection* - is there any?
2. Novelty *analysis* – what’s new, in what way, and to what degree?
3. Novelty *evaluation* - is that type of novelty significant?

Check for novelty against large database (e.g., WWW) of existing artifacts.

INTENTION:

May have to know the intention of the artist/program before you know how to evaluate.

For a program, the intention is probably due to the programmer.

For a program, any evaluation that is included is probably due to the programmer.

In-program evaluation forms the basis for an aesthetic judgment.

Alternatively could evaluate based on intention of viewer/consumer.

MATURITY:

It may take 10-12 years of intense immersion before transformational creativity occurs.

Evaluation goes hand-in-hand with the generation: i.e., it matures too.

Expertise plays a role in evaluation

- knowledge is necessary for high performance.

Knowledge compilation can make explicit evaluation become embedded in the generation process, producing artifacts that will be considered to have high value.

SUMMARY:

For visual artistic artifacts...

- evaluation is multidimensional.
 - evaluation includes context.
 - need to determine which attributes are significant.
 - e.g., personal ones (artist's? consumer's?)
 - cultural ones?
 - can we actually “extract” them from artists?
 - can we actually “extract” them from consumers?
 - can machine learning play a part in this?
 - *computational evaluation is complicated!*
-