

# Evaluation of CC systems

Ritchie, Colton, FACE & IDEA, SPECS, Perez

Some Empirical Criteria for Attributing Creativity  
to a Computer Program  
- **Ritchie 2007**

# Ritchie 2007

Introduce a system to evaluate creative ability of a CC system

# Assumptions - Ritchie 2007

- Creativity must be defined based on its **ordinary usage** that is, painting, music, prose writing etc
- We should consider only **empirically observable factors** and internal workings of a program are not part of the relevant data for evaluation

# Assumptions - Ritchie 2007

- **Human judgements** are used to evaluate the creativity of the program
- Essential properties - Typicality, Novelty, Quality
- P-creativity by computers are also considered creative.

**Note :** Humans judge only typicality and quality. Novelty will be affected since humans tend to value H-creativity more

# Essential framework - Ritchie

- Output should be of datatype that is close to original works. Eg. Verbal jokes can be a set of finite words and punctuations (not byte file)
- The program is influenced by *inspiring set (I)* which could be used to measure the creativity of the program
- The outputs need to be rated for class membership and quality of artefact by humans, novelty depending on learning data set

Note : Expert systems not considered creative(evaluated as exception)

# Evidence of creativity - Ritchie

1. Need to classify as belonging to a particular type of artefact by humans
2. Conform to established norms of the class (typicality)
3. Quality of the artefact must be rated greater than threshold by humans

# Evaluation of creativity - Ritchie 2007

4. Subset of artefact that scored high in typicality (2) and quality ( 3) is considered creative
5. Are large portions of artefacts generated by the program part of subset in 4



# Evaluation of creativity - Ritchie 2007

6. What portion of artefact which scored low on typicality (2) gained high ratings in quality (step 3)

# Evaluation of creativity - Ritchie 2007

7. Producing artefacts different from *I* is considered creative, with more than the threshold quality and typicality (typicality can be less for a few systems with high quality factor)

And more combinations of typicality, quality and novelty scores

Creativity Versus the Perception of Creativity in  
Computational Systems - Colton 2008

## Colton 2008

It is not appropriate to base the assessment of a system on its output alone. The way it produces artefacts also needs to be taken into account

Eg - NPR (Non-photorealistic rendering) vs painting fool



# Value of artefact - Colton 2008

- With time, the label of creative is bestowed upon the artist rather than his artefact
- Consumers make judgements based on (a) the effort behind the process (b) the ingenuity in devising the process, and (c) the skill required to undertake the process



# Value of artefact - Colton 2008

- Conceptual art is not about forms or materials, but about ideas and meanings
- Artists are expected to create both at the conceptual and the craft level
- Stigma that computers cannot be creative - Non-ambitious software (NPR) are more accepted than systems that are creative

# Creativity of CC systems - Colton 2008

**Creative Tripods** - Skillful, appreciative and imaginative. Essential characteristics for creative humans and programs

Each of the three legs can be contributed to by either the computer, programmer or the user

Eg.

AARON lacks appreciation (evaluation) - hence not creative

Computational Creativity Theory:  
The FACE and IDEA Descriptive Models - Colton,  
Charnley, Pease



# Computational Creativity Theory

- Explain in detail about two constituents of CCT - **FACE** descriptive model of creative acts as tuples of generative acts, and the **IDEA** descriptive model of the impact such creative acts may have

# FACE model - Levels of creative acts

(F)raming information and (A)esthetic considerations which can be used to assess (C)oncepts and (E)xamples of those concepts

**Ground level** - generative systems produce new artefacts such as theorems, pictures, compositions, poems, etc using existing processes

**Process level** - new ways to generate and assess artefacts are invented

# Face model - ground level

(Concept(g), Expression of Concept (g), Aesthetics (g), Framing (g))

Concept - Program that creates artefact

Expression of Concept - Does the program take input and give output

Aesthetics measure - Is there evaluation of the artifact. Takes a  
(Concept, Expression) and gives value between  $(0, \infty)$

Framing - Natural language explanation of some of the above (Providing cultural/Historical context, describe generative processes, etc)

# Face model -Process

(Concept (p), Expression of Concept(p), Aesthetics(p), Framing(p))

Concept - Program that creates artefact

Expression of Concept - Does the program take input and give output

Aesthetics measure - Is there evaluation of the artifact. Takes a  
(Concept,Exprtession) and gives value between  $(0, \infty)$

Framing - Natural language explanation of some of the above (Providing  
cultural/Historical context, describe generative processes, etc)

# FACE model - Types of generative acts

Concept(g), Expression of Concept (g), Aesthetics (g), Framing (g), Concept (p),  
Expression of Concept(p), Aesthetics(p), Framing(p)

To describe a system, pick all the ones that are relevant - Such a tuple is called a **creative act**

**Bar notations** are to highlight generative acts undertaken by a third party (usually humans).

# FACE model - Creative acts

A creative act is a non-empty tuple of generative acts

Eg.  $\langle F^g, A^g, C^g, E^g \rangle$  denotes a creative act comprising a 4-tuple of generative acts

Eg. , the tuple  $\langle \overline{A^g}, \overline{C^g}, E^g \rangle$  represents the creative act  $\overline{a^g}$ . is performed by the software whereby an expression of user-given concept  $\langle E^g \rangle$  was generated and assessed by user-given aesthetic measure  $\overline{c^g}$

# Evaluation using FACE model

1. **Quantitative way** - sheer volume of **creative acts** used to compare creative systems

# Evaluation using FACE model

2. **Cumulative way** - when you have a longer sequence of generative acts

Eg.

$$CA_1 = \langle C^g, E^g \rangle < CA_2 = \langle A^g, C^g, E^g \rangle$$



# Evaluation using FACE model

3. **Comparative way** - one type of generative act is seen as more creative than another

Eg.  $\langle C^P \rangle > \langle A^g \rangle > \langle C^g \rangle > \langle E^g \rangle$

4. **Process based way** - random generation can be seen as less creative than the inductive methods

# Evaluation using FACE model

5. **Qualitative way** - the values from aesthetic functions are used to compare creative acts