

Computational Creativity Autumn School II

Philosophical Issues

Simon Colton

Computational Creativity Group
Department of Computing
Goldsmiths College, University of London

ccg.doc.gold.ac.uk

Aims

- This morning:
 - To raise and discuss some of the issues that affect us all in Computational Creativity research
- This afternoon:
 - To address these philosophical issues with some practical guidance on how to build and assess creative systems

Computational Creativity...

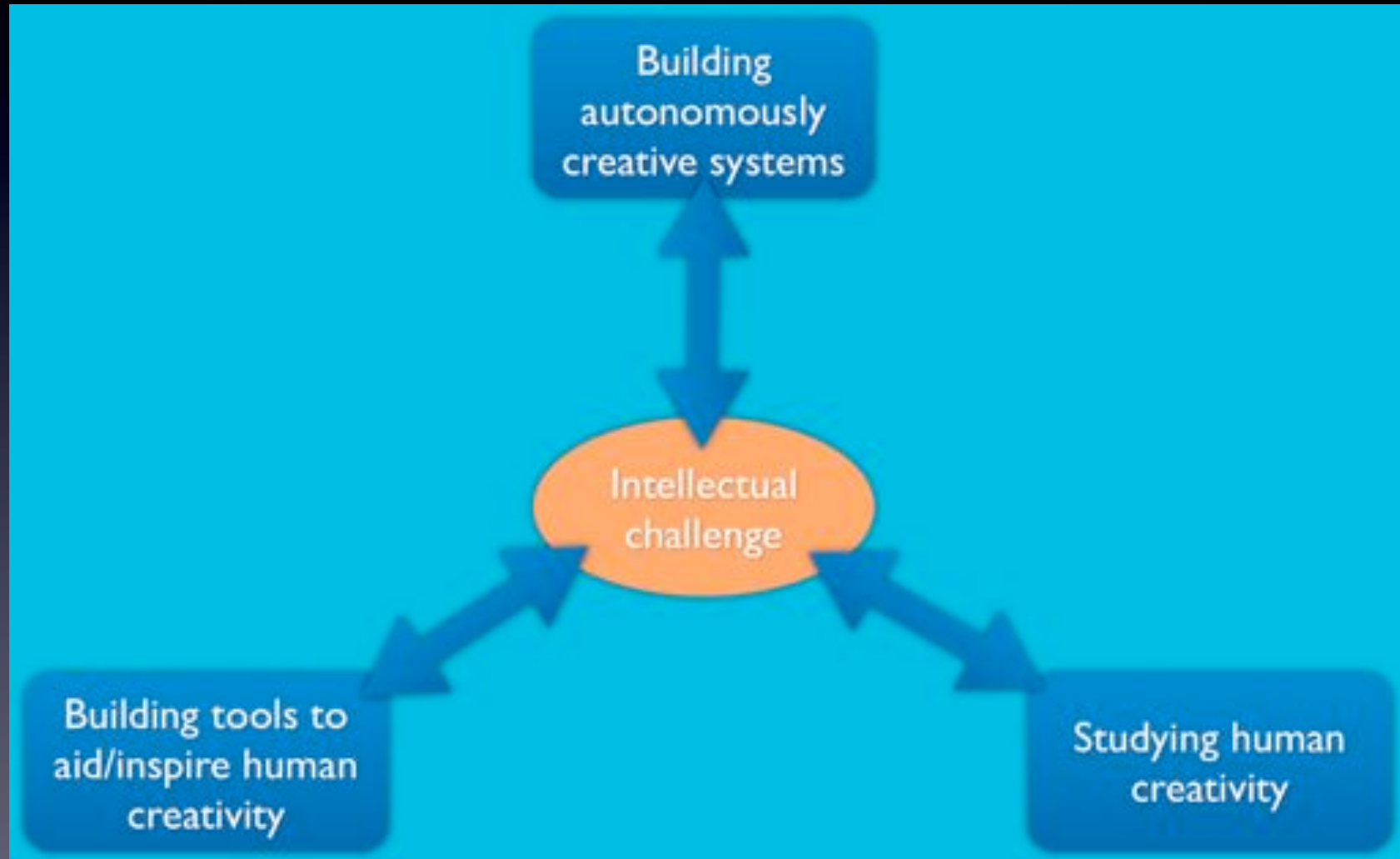
Creative
responsibilities

The philosophy, science and engineering of computational systems which, by taking on particular responsibilities, exhibit behaviours that unbiased observers would deem to be creative.

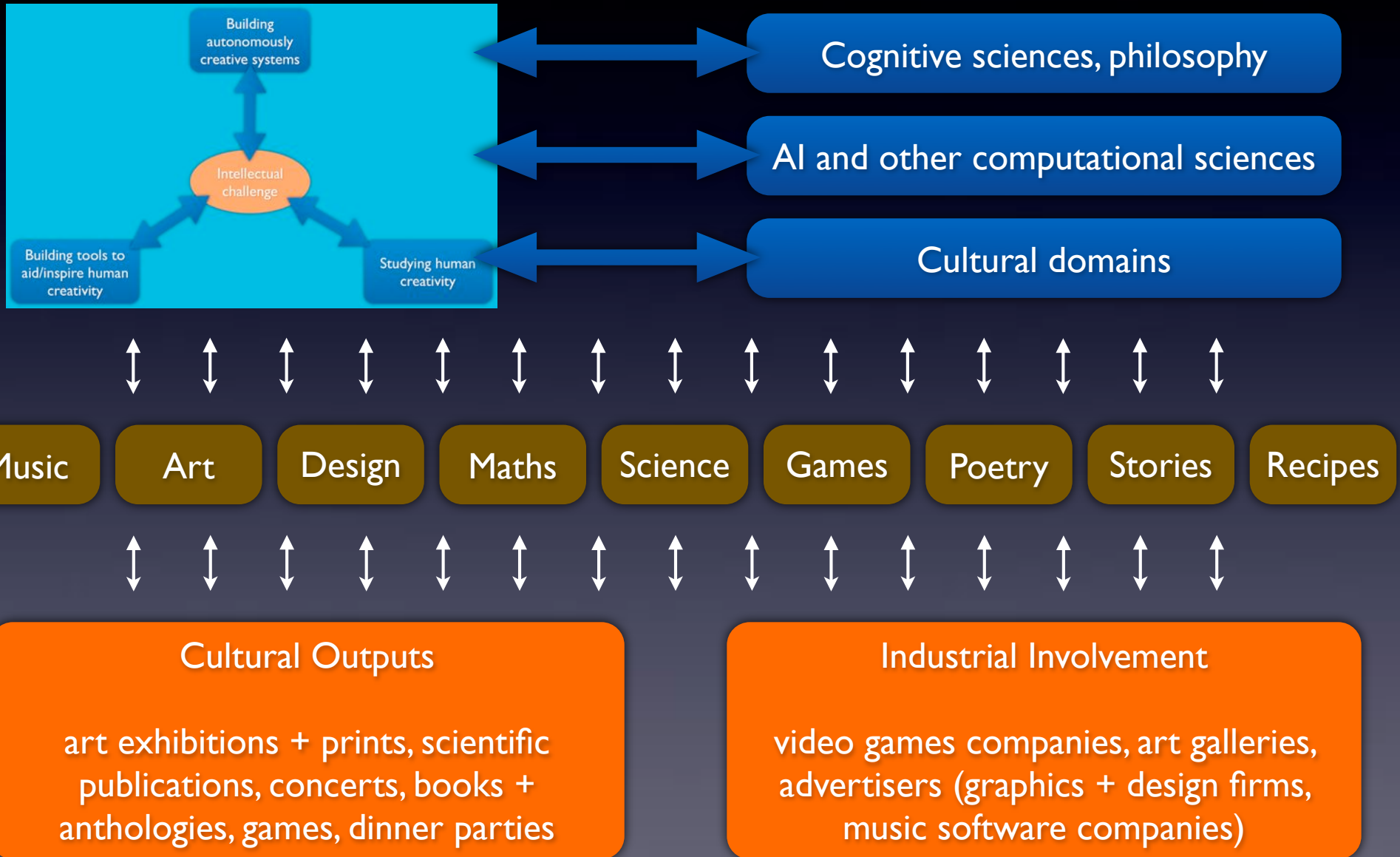
Audience
participation

Also note the deliberate lack of mention of *value* of generated artefacts (poems, paintings, theorems, etc.) and the lack of mention of comparison with *people*

Overview



Overview



Issues in the Field

- With the word 'creative'
- With handing over creative responsibilities
- With evaluating software which creates
- With software not being human

Some Difficult Notions to Digest...

There's no such thing as creativity

We shouldn't agree on how people perceive creativity

Psychology envy can be a bad thing

Levelling the playing field can go badly wrong

Output quality and autonomy of software
can be inversely proportional

We don't all agree in Computational Creativity!

I. Our issues with the word 'creative'





WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction

Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

Print/export

Languages

Deutsch
Français
Nederlands

[Edit links](#)

[Create account](#) [Log in](#)

Article [Talk](#)

[Read](#) [Edit](#) [View history](#)

Essentially contested concept

From Wikipedia, the free encyclopedia

In a paper delivered to the *Aristotelian Society* on 12 March 1956,^[1] [Walter Bryce Gallie](#) (1912–1998) introduced the term **essentially contested concept** to facilitate an understanding of the different applications or interpretations of the sorts of abstract, qualitative, and evaluative notions^[2] — such as "art" and "social justice" — used in the domains of aesthetics, political philosophy, philosophy of history, and philosophy of religion.

Garver (1978) describes their use as follows:

The term essentially contested concepts gives a name to a problematic situation that many people recognize: that in certain kinds of talk there is a variety of meanings employed for key terms in an argument, and there is a feeling that *dogmatism* ("My answer is right and all others are wrong"), *skepticism* ("All answers are equally true (or false); everyone has a right to his own truth"), and *eclecticism* ("Each meaning gives a partial view so the more meanings the better") are none of them the appropriate attitude towards that variety of meanings.^[3]

Essentially contested concepts involve widespread agreement on a concept (e.g., "fairness"), but not on the best realization thereof.^[4]

They are "concepts the proper use of which inevitably involves endless disputes about their proper uses on the part of their users",^[5] and these disputes "cannot be settled by appeal to empirical evidence, linguistic usage, or the canons of logic alone".^[6]

Contents (hide)

- Identifying the presence of a dispute
 - Contested versus contestable?
- Features
- Concepts and conceptions
- Not "hotly disputed" concepts
- See also
- Notes
- References

References

Notes

See also

Essentially contested concept

From Wikipedia, the free encyclopedia

In a paper delivered to the *Aristotelian Society* on 12 March 1956,^[1] **Walter Bryce Gallie** (1912–1998) introduced the term **essentially contested concept** to facilitate an understanding of the different applications or interpretations of the sorts of abstract, *qualitative*, and *evaluative* notions^[2] — such as "art" and "social justice" — used in the domains of *aesthetics*, *political philosophy*, *philosophy of history*, and *philosophy of religion*.

Garver (1978) describes their use as follows:

The term essentially contested concepts gives a name to a problematic situation that many people recognize: that in certain kinds of talk there is a variety of meanings employed for key terms in an argument, and there is a feeling that *dogmatism* ("My answer is right and all others are wrong"), *skepticism* ("All answers are equally true (or false); everyone has a right to his own truth"), and *eclecticism* ("Each meaning gives a partial view so the more meanings the better") are none of them the appropriate attitude towards that variety of meanings.^[3]

Essentially contested concepts involve widespread agreement on a concept (e.g., "fairness"), but not on the best realization thereof.^[4]

They are "concepts the proper use of which inevitably involves endless disputes about their proper uses on the part of their users",^[5] and these disputes "cannot be settled by appeal to *empirical evidence*, *linguistic usage*, or the *canons of logic alone*".^[6]

Contents [hide]

- Identifying the presence of a dispute
 - Contested versus contestable?
- Features
- Concepts and conceptions
- Not "holy disputed" concepts
- See also
- Notes
- References

Credit to Anna Jordanous for highlighting creativity in this context

That Word...

- ‘Creative’ is just a word that one person uses to describe another person, like ‘funny’. It’s a word to describe a *perception* people have, not an inherent property of someone
- If everyone agrees that they perceive someone as being creative, then it’s fair to call them creative, not just ‘perceived to be creative’
- It can also be used to describe a process/behaviour... often shorthand for the person
- Artefacts are generally not ‘creative’. In common parlance, a ‘creative building’ or a ‘creative metaphor’ can mean one of a number of different things
- As scientists, we should be more precise in our usage of this word
- We might disagree between ourselves, but we should be internally coherent
- If we ask vague questions about the “creativity” of a person, process, or worse, the “creativity” of a building or poem, we should expect to learn zero from the study, except (yet again) that people use the word in different ways
- We have plenty of other words we could use to help people assess quality of output, e.g., beautiful. And we have plenty of other words to use in domains, e.g., artistic

In Computational Creativity research, when talking about software, we should only use the word 'creative' to describe how people perceive what our software actually does

And we shouldn't ask that question directly

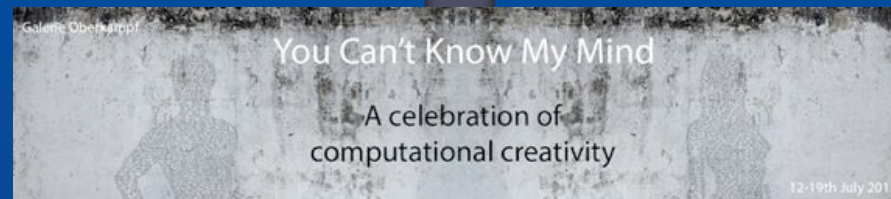
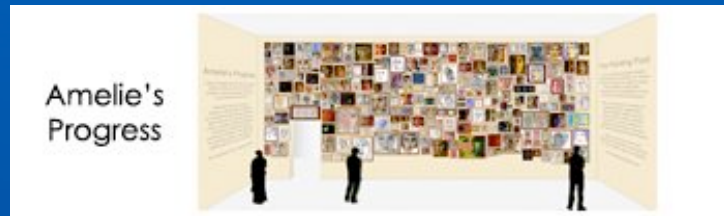
2. Our issues with handing over creative responsibilities

Weak and Strong Computational Creativity Subprojects

- Weak subprojects
 - “I want software to create wonderful artefacts of type X”
 - “I’m more interested in the domain of X, and how we can contribute to that culture than to simulating creativity”
- Strong subprojects
 - “I want to build software which is one day taken seriously as being creative in generating wonderful artefacts of type X”
 - “I’m more interested in the study of creativity in software, and I want to use domain X to further study that”

Some Painting Fool Subprojects

Weak



Strong



The Painting Fool

You Can't Know my Mind
www.thepaintingfool.com

I was in a negative mood.
So I wanted to paint a bleached portrait.
I aimed to achieve something like this:



And this is my painting:



Overall, this portrait is not bleached at all.
And worse, my style has significantly lowered the level of bleached here.
So this is a miserable failure - I'm very unhappy about that.



Being Seen to Be AI

Automated Poetry Generation

- Strongly creative software cannot just produce valuable artefacts (poems, sonatas, theorems, paintings)
 - It has to do so in intelligent and (relatively) difficult to follow ways (not necessarily randomly)
 - And it needs to convince audiences that it has behaved in interesting and creative ways
- Will hopefully turn a vicious circle into a virtuous circle
- Practical implications:
 - Software should produce commentaries, then stories and ultimately be able to answer questions...

Circadian No. 39

_____ of a _____,
_____ of a _____.

In the morning, I am _____
Like the _____ of a _____,
But the morning grows more _____
Than the _____ of a _____,
And the mid-day makes me _____
Like the _____ of a _____.

Active _____ of a _____,
_____ of a _____.

In the daytime, I am _____
Like the _____ of a _____,
But the sunlight grows more _____
Than the _____ of a _____,
And the day, it makes me _____
Like the _____ of a _____.

_____ of a _____,
_____ of a _____.

In the _____, I am _____
Like the _____ of a _____,
But the evening grows more _____
Than the _____ of a _____,
And the _____ makes me _____
Like the _____ of a _____.

_____ of a _____,
_____ of a _____.

In the night-time, I am _____
Like the _____ of a _____,
But the moonlight grows more _____
Than the _____ of an _____,
And the darkness makes me _____
Like the _____ of a _____.

_____ of a _____,
_____ of a _____.

Circadian No. 39

Stealthy swiftness of a leopard,
Happy singing of a bird.

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

Active frenzy of a beehive,
Dreary blackness of a cave.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.

Shiny luster of a diamond,
Homey feeling of a bed.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Famous fervor of a poet,
Wily movement of a cat.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Circadian No. 39

Stealthy swiftness of a leopard,
Happy singing of a bird.

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

Active frenzy of a beehive,
Dreary blackness of a cave.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.

Shiny luster of a diamond,
Homey feeling of a bed.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Famous fervor of a poet,
Wily movement of a cat.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Repetition

Frequency

Stemming

Metre

Rhyme

Non-rhyming

Sentiment

Advertising

Circadian No. 39

Stealthy swiftness of a leopard,
Happy singing of a bird.

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

Active frenzy of a beehive,
Dreary blackness of a cave.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.

Shiny luster of a diamond,
Homey feeling of a bed.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Famous fervor of a poet,
Wily movement of a cat.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Repetition

Frequency

Stemming

Metre

Rhyme

Non-rhyming

Sentiment

Advertising

Circadian No. 39

Stealthy swiftness of a leopard,
Happy singing of a bird.

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

Active frenzy of a beehive,
Dreary blackness of a cave.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.

Shiny luster of a diamond,
Homey feeling of a bed.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Famous fervor of a poet,
Wily movement of a cat.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Repetition

Frequency

Stemming

Metre

Rhyme

Non-rhyming

Sentiment

Advertising



Circadian No. 39

da da da da di di da da

Stealthy swiftness of a leopard,
Happy singing of a bird.

da da da da di di daaa

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

Active frenzy of a beehive,
Dreary blackness of a cave.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.

Shiny luster of a diamond,
Homey feeling of a bed.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Famous fervor of a poet,
Wily movement of a cat.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Repetition

Frequency

Stemming

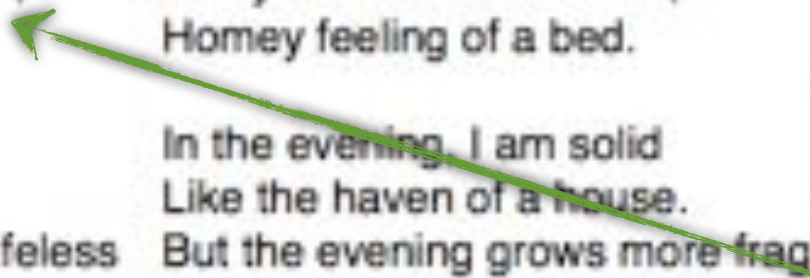
Metre

Rhyme

Non-rhyming

Sentiment

Advertising



Circadian No. 39

Stealthy swiftness of a leopard,
Happy singing of a bird.

Shiny luster of a diamond,
Homey feeling of a bed.

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Active frenzy of a beehive,
Dreary blackness of a cave.

Famous fervor of a poet,
Wily movement of a cat.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Repetition

Frequency

Stemming

Metre

Rhyme

Non-rhyming

Sentiment

Advertising

Circadian No. 39

Stealthy swiftness of a leopard,
Happy singing of a bird.

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

Active frenzy of a beehive,
Dreary blackness of a cave.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.

Shiny luster of a diamond,
Homey feeling of a bed.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Famous fervor of a poet,
Wily movement of a cat.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Repetition

Frequency

Stemming

Metre

Rhyme

Non-rhyming

Sentiment

Advertising

Circadian No. 39

Stealthy swiftness of a leopard
Happy singing of a bird.



Shiny luster of a diamond,
Homey feeling of a bed.



In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.



In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.



Active frenzy of a beehive,
Dreary blackness of a cave.



Famous fervor of a poet,
Wily movement of a cat.



In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comely
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke.



In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more supple
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.



Stealthy swiftness of a leopard
Happy singing of a bird.



Repetition

Frequency

Stemming

Metre

Rhyme

Non-rhyming

Sentiment

Advertising

Circadian No. 39

Stealthy swiftness of a leopard,
Happy singing of a bird.

In the morning, I am loyal
Like the comfort of a friend.
But the morning grows more lifeless
Than the fabric of a rag.
And the mid-day makes me nervous
Like the spirit of a bride.

Active frenzy of a beehive,
Dreary blackness of a cave.

In the daytime, I am slimy
Like the motion of a snake.
But the sunlight grows more comfy
Than the confines of a couch.
And the day, it makes me tasty
Like the flavor of a coke. ←

Shiny luster of a diamond,
Homey feeling of a bed.

In the evening, I am solid
Like the haven of a house.
But the evening grows more fragile
Than the mindset of a child.
And the twilight makes me frozen
Like the bosom of a corpse.

Famous fervor of a poet,
Wily movement of a cat.

In the night-time, I am hollow
Like the body of a drum.
But the moonlight grows more su
Than the coating of an eel.
And the darkness makes me subtle
Like the color of a gem.

Stealthy swiftness of a leopard,
Happy singing of a bird.

Repetition

Frequency

Stemming

Metre

Rhyme

Non-rhyming

Sentiment

Advertising

Searching Questions #1

My point is:

- [There are possibly benefits (in terms of increased value projected onto the artefacts) to be gained by presenting explicit information about the process behind the generation of an artefact
- [We can provide this information through technical papers, talks and notes, but it would be better for the software to do this itself, because it simulates appreciation and/or reflection, which we value in creative people

ease

ving

about these constraints, and throw in some more stuff?

Searching Questions #2

My point is:

There may be a virtuous circle (rather than the usual vicious circle) we can get into, if we know that the software has done something difficult and intelligent.

People will always read/see/hear things that aren't in artefacts by design. If those people know a little about the intelligence of the system, then they might be prepared to give software the benefit of the doubt

Software could make stuff up (tell a story...)

Searching Questions #3

My point is:

- This is a clearly a Simon-generated poem
- This doesn't feel like Computational Creativity research yet
- The software needs to take on more creative responsibility in the production of the poems...

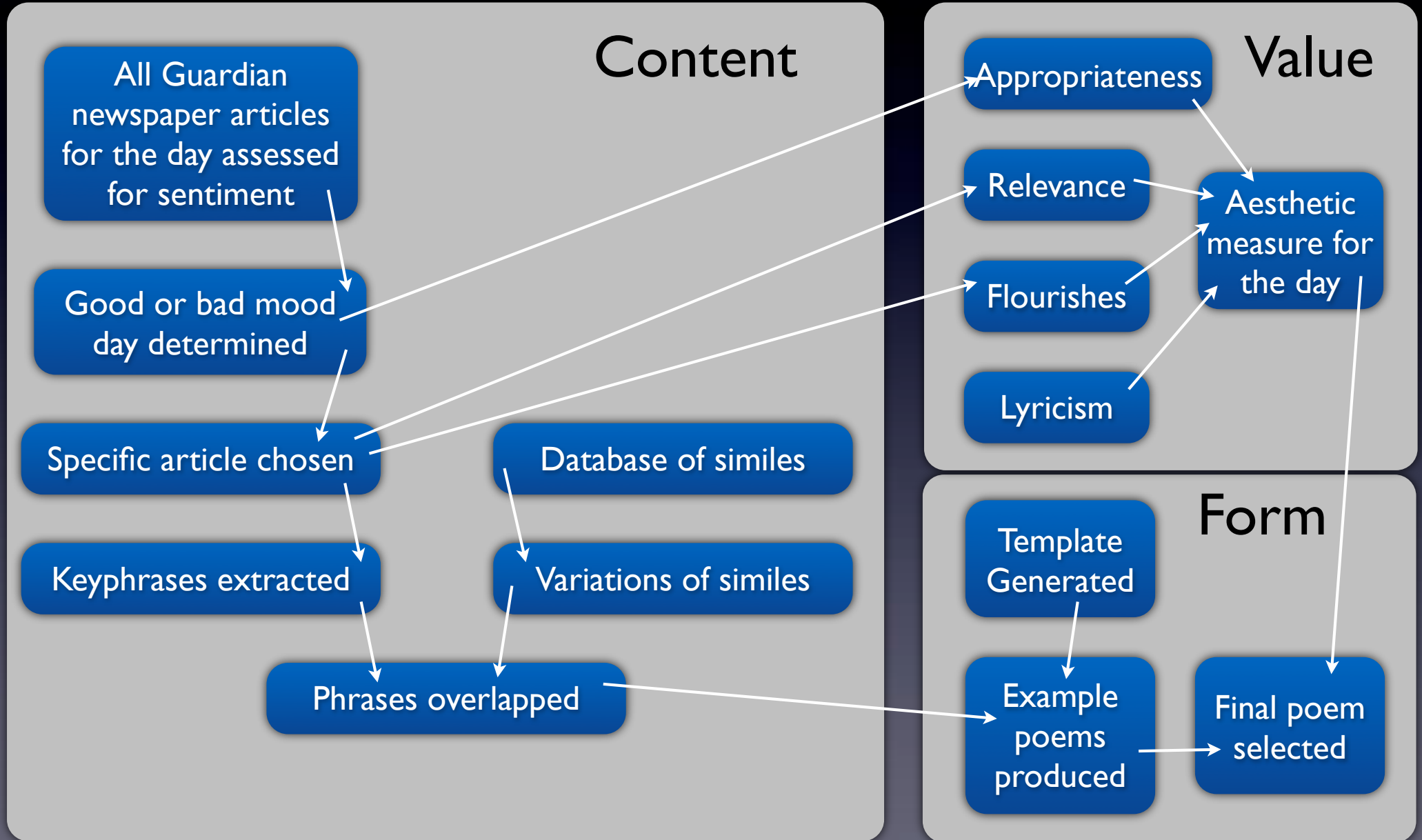
Handing over Creative Responsibility

The FACE Model

- We can judge progress in terms of the types of generative acts that software undertakes (regardless of the value of its output)
- Examples - simile multiplication, phrase overlapping
- Concepts - generating templates
- Aesthetics - inventing measures of value
- Framing - producing a commentary

Handing over Creative Responsibility

Poetry Generation Pipeline



New Poems...

It was generally a good news day. I read a story in the Guardian culture section entitled: "South Africa's ANC celebrates centenary with moment in the sun". It talked of south africans, interfaith prayers and monochrome photos. Apparently, "The heroic struggle against a racist regime was remembered: those thousands who sacrificed their lives in a quest for human rights and democracy that took more than eight decades" and "At midnight he watched with amusement as Zuma lit the centenary flame, at the second attempt, with some help from a man in blue overalls marked 'Explosives"'. I wanted to write something highly relevant to the original article. I wrote this poem.

Blue overalls

the repetitive attention of some traditional african chants
a heroic struggle, like the personality of a soldier

an unbearable symbolic timing, like a scream
blue overalls, each like a blueberry
some presidential many selfless leaders

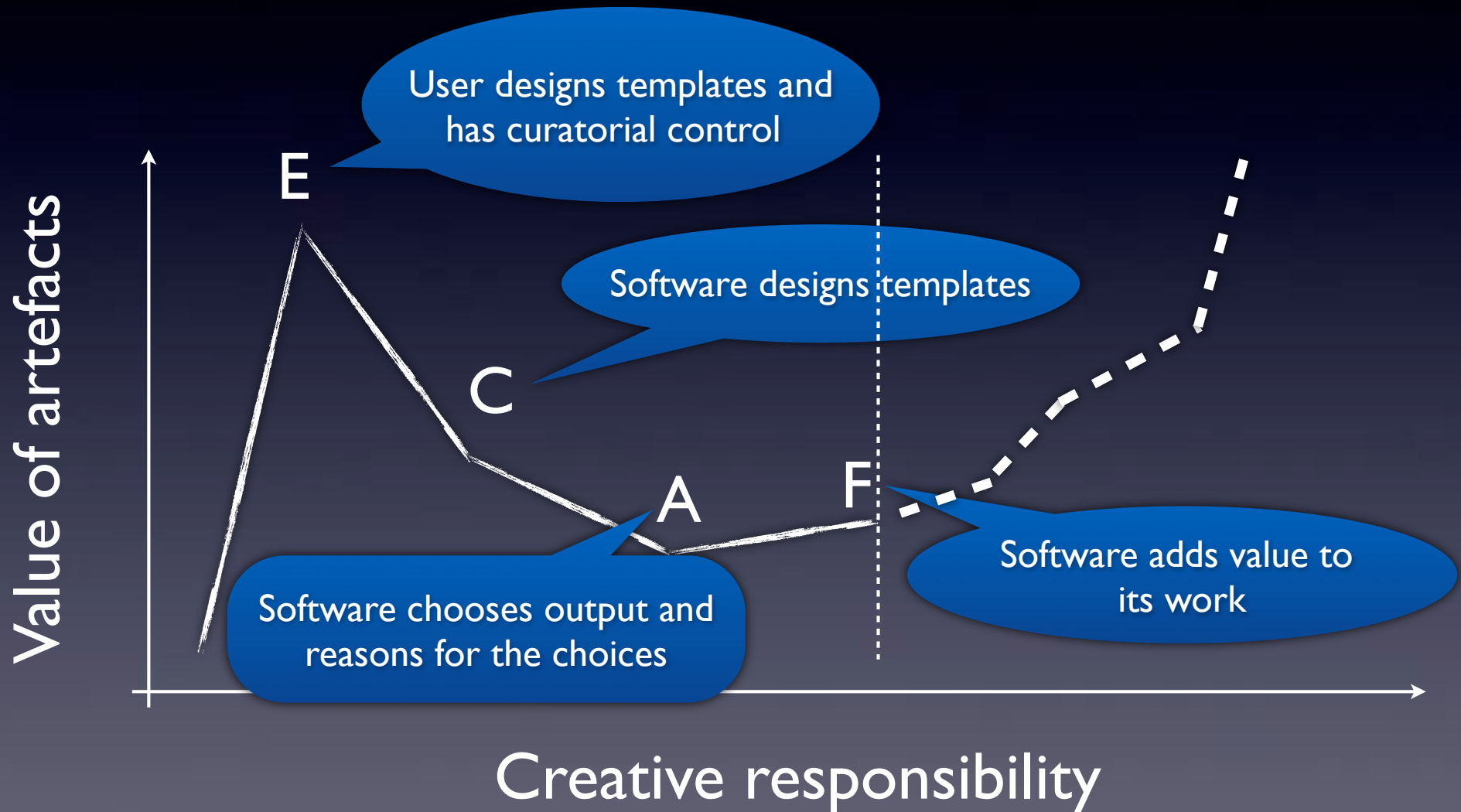
oh! such influential presidents
such great presidents
blueberry-blue overalls

lark-blue overalls
a knight-heroic struggle

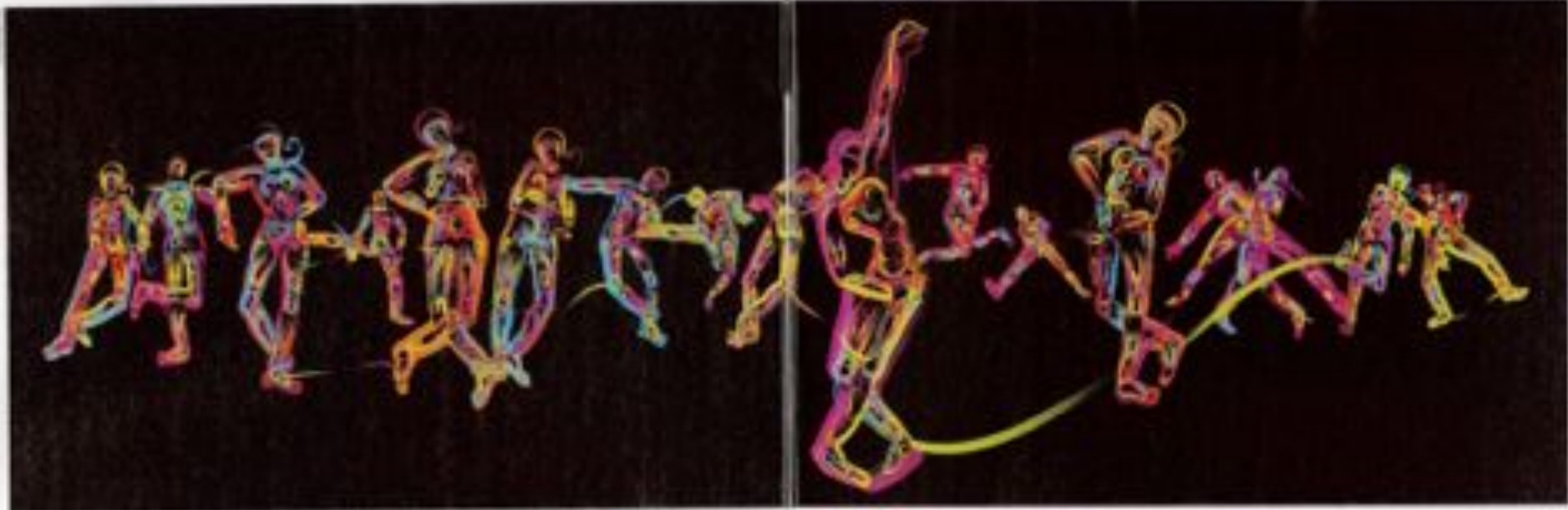
Discussion...

- Taken out of context, e.g., in an evaluation test or a Turing-style test, poem #1 would probably score higher as a “poem shaped object” than poem #2
- But, when we read about *how* the software produced the poems, it’s likely that people will project more creativity onto the software producing poem #2
- So, the more sophisticated software represents a backward step in automated poetry generation, but an advance in computational creativity
- The Latent Heat effect in Computational Creativity

The Latent Heat Effect in Computational Creativity



3. Our issues with
evaluating software
which creates



Creative sparks

Does it really take a human to make a masterpiece? Catherine de Lange sizes up the artificial artists forcing us to change our ideas about creativity

In a world including the works of some of the greatest artists, might the works of artificial artists be more creative than those of their human counterparts? In other words, would a computer-generated painting be more creative than a painting by a human? It's a question that has been asked for decades, but it's one that's become more relevant than ever in our digital age.

The idea of a computer-generated painting is not new. In fact, it's been around since the early days of digital art. But it's only in recent years that we've seen the rise of artificial artists like the ones we're talking about here.

These artificial artists are not just simple programs that generate random patterns. They are sophisticated systems that can learn from examples and create new, original works of art. They can even be trained to create art in a specific style, like the Impressionist or the Cubist.

The technology is so advanced now

that it can generate works that are indistinguishable from those of human artists. In fact, some of the most impressive works of art created by artificial artists have been sold for millions of dollars. This is a testament to the power of artificial intelligence and the potential of artificial artists.

But what does this mean for the future of art? Will artificial artists eventually replace human artists? Or will they simply become a new tool for human artists to use? These are the questions that we need to ask as we move forward into a world where artificial intelligence is becoming increasingly prevalent.

As we move forward, we need to think about the role of art in our lives. Art is not just a hobby or a profession; it's a way of seeing the world and expressing our thoughts and feelings. If we let artificial artists take over, we risk losing the human touch that makes art so special.

Art is a reflection of the human condition, and it's something that we need in our lives. We need to see our own struggles and triumphs reflected in art. We need to see the beauty and the pain of the world as we know it. And we need to see it all in a way that is both beautiful and honest.

So, while it's true that artificial artists can create beautiful works of art, they can't create art that is truly meaningful. They can't create art that speaks to the human experience in a way that only a human can. And that's why we need to keep supporting human artists and the art that they create.

Art is a reflection of the human condition, and it's something that we need in our lives.

"The machine will make up in the morning and look at newspaper headlines for source material"

Depressed and somewhat a computer... (the machine will make up in the morning and look at newspaper headlines for source material)

will be less straightforward to create a masterpiece. Although it's software, it's not just a machine; it's a machine that's been trained to create art. And that's what makes it so interesting. It's a machine that's been trained to create art in a way that is both beautiful and honest.

The machine will make up in the morning and look at newspaper headlines for source material. This is a statement that's both funny and serious. It's a statement that's both a warning and a challenge. It's a statement that's both a reflection of the current state of the world and a prediction of the future.

It's a statement that's both a warning and a challenge. It's a statement that's both a reflection of the current state of the world and a prediction of the future.

...the machine will make up in the morning and look at newspaper headlines for source material

...the machine will make up in the morning and look at newspaper headlines for source material

2017C6 W/9/6U9L
UG6M2b9b6L μ6991UG2 10L

...the machine will make up in the morning and look at newspaper headlines for source material

professor to write the review of the book in a magazine, based on a newspaper. The result is a scathing critique of the book's claims, arguments and evidence. "The piece itself is a masterpiece, and it shows the potential for the authors to add judgments and commentary to their findings," says Cohen.

The book's impact is also beginning to display itself in a number of interesting and new ways. One of its original works, part of a collection called *Color by Numbers*, depicts a variety of people in a landscape. The image below is a detail of one of the people in the collection.

Comparable to Bach

Starting with that, how meaningful is it that a computer-generated portrait can be produced and then produced in a similar way, especially when the painting is a portrait of a person? "It's a bit of a stretch to say that it's a portrait of a person," says Cohen. "The work is a portrait of a person, but it's not a portrait of a person. It's a portrait of a person, but it's not a portrait of a person."

Another thing that is interesting is that the computer-generated portrait can be produced and then produced in a similar way, especially when the painting is a portrait of a person. "It's a bit of a stretch to say that it's a portrait of a person," says Cohen. "The work is a portrait of a person, but it's not a portrait of a person."

Another thing that is interesting is that the computer-generated portrait can be produced and then produced in a similar way, especially when the painting is a portrait of a person. "It's a bit of a stretch to say that it's a portrait of a person," says Cohen. "The work is a portrait of a person, but it's not a portrait of a person."



Four different portraits of a person, each with a different color scheme.

"After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him."

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

His response was simple, however: "I'm not a musician, but I'm a composer. I'm not a musician, but I'm a composer. I'm not a musician, but I'm a composer."

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

The portrait of a person, each with a different color scheme.



...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

Unconscious Creativity

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

...of his music, who 'killed music' and tried to punch him. After discovering the truth, one music lover told Cope he had 'killed music' and tried to punch him.

A Turing Test for Being Creative

- Probably because of our longing for objectivity, it's easy to mix up the process of *generating high quality artefacts*, with *being creative*
 - Pastiches can be high quality, yet it's unlikely that you would call the artist/writer/musician particularly creative, as there's no new dialogue arising from it
- You might argue: "What does it matter, as long as high quality output is produced?"
 - Until you realise that as a society, we really value our creative individuals - perhaps more than what they produce
 - And that when we celebrate an artefact (poem/picture/sonata), we are actually (also) celebrating the creative act that led to it
 - And for strong Computational Creativity, creative behaviour is paramount

On impact and evaluation in Computational Creativity: A discussion of the Turing Test and an alternative proposal

Alison Pease¹ and Simon Colton²

Abstract. Computational Creativity is the AI subfield in which we study how to build computational models of creative thought in science and the arts. From an engineering perspective, it is desirable to have concrete measures for assessing the progress made from one version of a program to another, or for comparing and contrasting different software systems for the same creative task. We describe the Turing Test and versions of it which have been used in order to measure progress in Computational Creativity. We show that the versions proposed thus far lack the important aspect of interaction, without which much of the power of the Turing Test is lost. We argue that the Turing Test is largely inappropriate for the purposes of evaluation in Computational Creativity, since it attempts to homogenise creativity into a single (human) style, does not take into account the importance of background and contextual information for a creative act, encourages superficial, uninteresting advances in front-ends, and rewards creativity which adheres to a certain style over that which creates something which is genuinely novel. We further argue that although there may be some place for Turing-style tests for Computational Creativity at some point in the future, it is currently unworkable to apply any defensible version of the Turing Test.

As an alternative to Turing-style tests, we introduce two descriptive models for evaluating creative software, the FACE model which describes creative acts performed by software in terms of tuples of generative acts, and the IDEA model which describes how such creative acts can have an impact upon an ideal audience, given ideal information about background knowledge and the software development process. While these models require further study and elaboration, we believe that they can be usefully applied to current systems as well as guiding further development of creative systems.

1 The Turing Test and Computational Creativity

The Turing Test (TT), in which a computer and human are interrogated, with the computer considered intelligent if the human interrogator is unable to distinguish between them, is originally a philo-

Computational Creativity (CC) is a subfield of AI in which researchers aim to model creative thought by building programs which can produce ideas and artefacts which are novel, surprising and valuable, either autonomously or in conjunction with humans. There are three main motivations for the study of Computational Creativity:

- to provide a computational perspective on human creativity, in order to help us to understand it (cognitive science);
- to enable machines to be creative, in order to enhance our lives in some way (engineering); and
- to produce tools which enhance human creativity (aids for creative individuals).

Creativity can be subdivided into everyday problem-solving, and the sort of creativity reserved for the truly great, in which a problem is solved or an object created that has a major impact on other people. These are respectively known as "little-c" (mundane) and "big-C" (eminent) creativity [2]. Boden [3] draws a similar distinction in her view of creativity as search within a conceptual space, where "exploratory creativity" searches within the space, and "transformational creativity" involves expanding the space by breaking one or more of the defining characteristics and creating a new conceptual space. Boden sees transformational creativity as more surprising, since, according to the defining rules of the conceptual space, ideas within this space could not have been found before.

There are two notions of evaluation in CC: (i) judgements which determine whether an idea or artefact is valuable or not (an essential criterion for creativity) – these judgements may be made internally by whoever produced the idea, or externally, by someone else and (ii) judgements to determine whether a system is acting creatively or not. In the following discussion, by evaluation, we mean the latter judgement. Finding measures of evaluation of CC is an active area of research, both influenced by, and influencing, practical and theoretical aspects of CC. It is a particularly important area, since such measures suggest ways of defining progress in the field,¹ as well as

Turing-Style Tests

- Style 1: A dialogue where the point of the exercise is to prove that it would be fair to call your software intelligent
 - Closest to what Turing had in mind
- Style 2: A dialogue where the point of the exercise is to prove that people can't tell the difference to talking to a person and talking to your software
 - So, we implement software which often says unintelligent things
- Style 3: A comparison test with no dialogue, where the point of the exercise is to prove that the output of your software is of a similar (or higher) value to that produced by people
 - This has often been applied in Computational Creativity research

Comparison Tests

- It is certainly a milestone in the development of generative software (and for the field as a whole) if the output can be easily confused with that of people. This is because we can refer to the default position that people act creatively when they produce, and hence it is only fair to describe software similarly
- And it allows objective comparison, enabling us to show progress in implementations. Importantly, we can be seen to be scientific in our evaluation methodology
- And journalists love setting up Turing-style tests, as it both informs and worries the general public, which helps to sell newspapers...
 - New Scientist and BBC Horizon

However...

- Imagine a comparison test where the tester performs the *reveal*:
- “So, these paintings were painted by recent Royal College of Art graduates”
- “And these ones were painted by.....
a mass murderer!”
- Wouldn't your value judgements change?

Problems 1 and 2

- Turing-style comparison tests set the computer up for a fall
 - The implicit assumption is that software should be very grateful if it is mistaken occasionally for a human
 - So, human level output becomes seen as the only goal of Computational Creativity research
- Software is NOT human!
- So, we end up missing out on possibilities where the software creates valuable, interesting artefacts in non-human ways
- We should instead be loud and proud about the generative system being computer based, and help people to appreciate the value of computer generated creative acts

Problems 3 and 4

- Turing-style comparison tests massively underestimate the importance of process in certain domains
 - This can lead to alienation of people, certainly in the visual art world, where art theory is all about process
- Turing-style comparison tests answer the wrong question, e.g., which would you prefer, if you had to make up your mind without knowing fully how they were produced
 - Whereas in (commercial/artistic/scientific) reality, we will have full/partial disclosure of practice as well as product
 - Or should we go through this charade with our software for the rest of our lives?

Problems 5 and 6

- There are no right or wrongs in the visual arts. However, critics can severely inflict pain by saying that your work is “naive” and/or a “pastiche”
- Turing-style comparison tests might encourage software to act unintelligently, to make it seem more human, hence it could be criticised as naive
- Turing-style comparison tests definitely encourage the generation of pastiche pieces, as the measure of success is whether you have successfully imitated something which isn't you
- Would art graduates be happy if you said their pieces all looked like Monet pictures?

Well put by Alison...

of novelty and usefulness is accepted as key (for instance, see [1] or [3]). In [4], Plucker and Makel list “similar, overlapping and possibly synonymous terms for creativity: imagination, ingenuity, innovation, inspiration, inventiveness, muse, novelty, originality, serendipity, talent and unique”. The term ‘imitation’ is simply antipodal to many of these terms.

In the following sections we firstly describe and discuss some of

- Turing-style comparison tests are inappropriate for testing aspects of creative intelligence in software
- See paper for other arguments

Boden's

“A Turing Test for Artistic Creativity”

In [11], Boden discusses the Turing Test and artistic creativity. She provides an interpretation of the Turing Test which is specifically designed for computer art systems:

“I will take it that for an ‘artistic’ program to pass the TT would be for it to produce artwork which was:

1. indistinguishable from one produced by a human being; and/or
2. was seen as having as much aesthetic value as one produced by a human being.” [11, p. 409]

467–491. BASIC BOOKS, 1995. Epilogue in [25].

[11] M. A. Boden. The Turing test and artistic creativity. *Kybernetes*, 39(3):409–413, 2010.

[12] P. McCorduck. *Artificial Intelligence: A Guide for Humans*. Prentice-Hall, 1971.

Boden and Edmund's “Turing Test for Artistic Creativity”

Boden describes several systems which produce art or music, which she considers to be either non-interactive or unpredictably interactive (such as a piece of art which responds to audience members or participants in ways they do not understand). She discusses comparisons with both mediocre human art, in this case pastiches of given styles (perhaps comparable to work by an art student exploring a given style), as well as examples which match world class human art, of interest as an artwork in itself (comparable to work done by a practising artist). She argues that the following systems all pass (her version of) the TT:

- Richard Brown's Starfish
- Harold Cohen's AARON
- Art by Boden and Edmunds
- David Cope's EMI

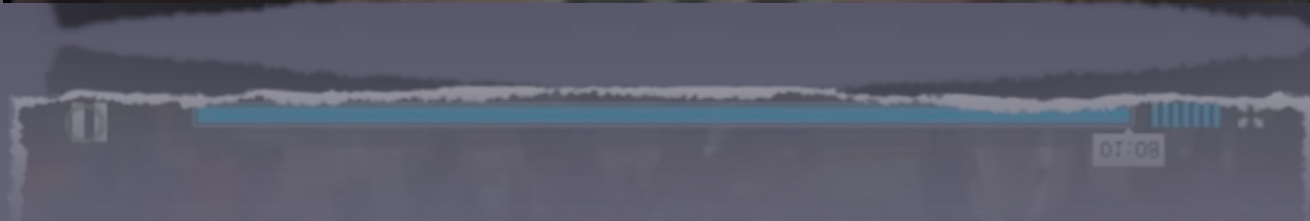
Boden's “Turing Test for Artistic Creativity”

In particular, Boden argues that “If being exhibited alongside Rothko, in a ‘diamond jubilee’ celebration of these famous artists, does not count as passing the Turing Test, then I do not know what would.” [11, p. 410].

Our Objections...

- It's an interpretation of Turing's test which bears little resemblance to the original idea
 - There is no dialogue or interaction of any kind with the system as part of the test
 - The test can be passed without comparison to human intelligence, or even human output
- So, it's possible to pass the originally conceived Turing test (testably achieving human-level intelligence), yet not pass Boden's test
 - Yet - as evidenced by the Starfish and by Boden and Edmunds' art - it's possible to pass Boden's test without exhibiting any higher level cognitive functions

The Starfish...



4. Our issues with software not being human

Human Level Creativity

- There are four main reasons to study human creativity with Computational Creativity research:
 - ✓ ● Because the artefacts produced are for human consumption (product)
 - ✓ ● It's ultimately human recognition of our software being creative that we seek (process)
 - ✓ ● People take computationally created artefacts and are creative with them (interpretation, etc)
 - ✗ ● We can program software to be (perceived to be) more creative by understanding well human creative processes

Psychology Envy

- Psychology research is great! And we should rightly be envious of their achievements
 - But its value to Computational Creativity research is not as high as you might imagine, and it's not a problem to realise this
- Problem 1: the results are often too vague to be turned into computational approaches, although we might get some general motivation or ideas
- Problem 2: because of the volume of knowledge about human creativity, it's too tempting to apply it to computational approaches to creativity
 - Software isn't human, e.g., it doesn't store and process information in the same way
 - It's often wholly inappropriate to analyse software in psychological terms, or pretend that software can be compared to people in meaningful ways
- Problem 3: there are some psychology experiments with methodologies that involve vaguely defined concepts. Our envy might lead us to overlook this and copy their flawed methodologies

Humanity Envy

- Treating our software as human blinds us to the fact that people's perception of software in society is hugely different to their perception of other people
- Me: "Should my software explain how it has produced a poem?"
- Tony: "...but people don't do that when they write poems!"
- Me: "But my software isn't a person"
- Tony: "It'll be seen as juvenile"
- Me: "But my software is juvenile!" ...
- Me later: "Hmmm. What does juvenile even mean when we are talking about software?"

I believe that, even if they have exactly the same words in exactly the same order, a computer generated poem should be seen as a *fundamentally different* type of artefact to a poem penned by a person

This is because the effect that each poem has on a reader is fundamentally different, due to the differences in how they were produced and what/who produced them

What do People Know about Human Poets?

They're like us
in many ways

They stay up all
night writing and
drinking coffee

They think hard
about their
poems

They live in Paris
and sell poems
for meals

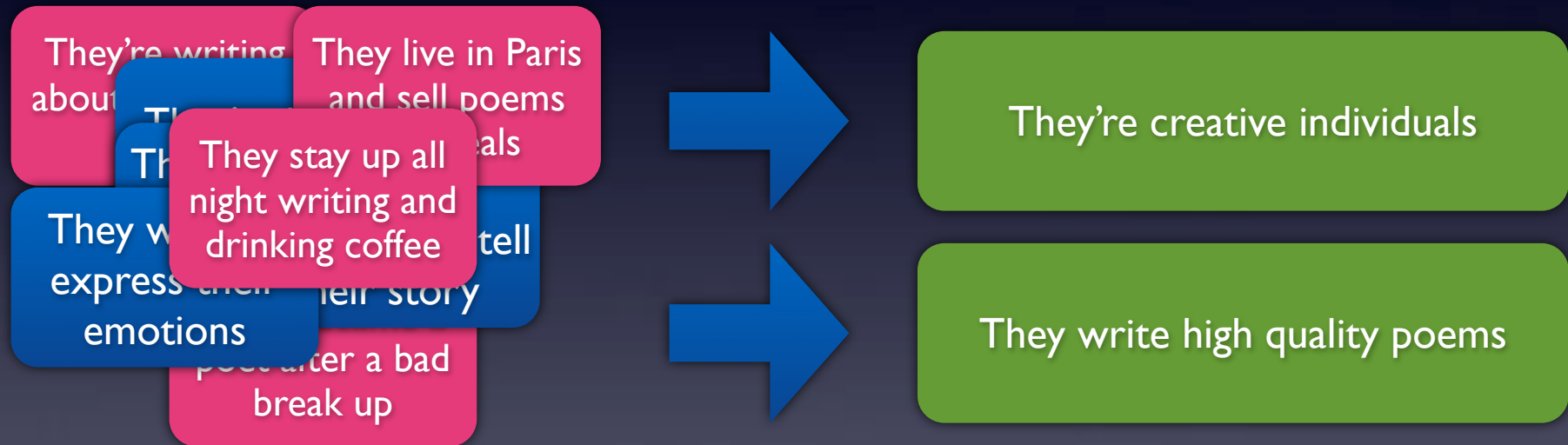
They're
motivated to tell
their story

They became a
poet after a bad
break up

They want to
express their
emotions

They're writing
about their dead
father

What do People Know about Human Poets?



What do People Know about ~~Human~~ Poets? Computer

They're dislike
us in sooo many
ways

They perform
trivial random
processing

They have no
story to tell

They are devoid
of personality

Rise of the Robot Artist

Greater than computers are good at processing information and humans are good at creating. What if it's the other way around?

By [Name] | [Date]



The painting above is an artwork created by a robot. It was generated by the artist program 'The Painting Machine' which is a computer program that generates abstract art. The program is based on the concept of 'The Painting Machine' which is a computer program that generates abstract art. The program is based on the concept of 'The Painting Machine' which is a computer program that generates abstract art.

Some Difficult Notions to Digest...

There's no such thing as creativity

We shouldn't agree on how people perceive creativity

Psychology envy can be a bad thing

Levelling the playing field can go badly wrong

Output quality and autonomy of software
can be inversely proportional

We still don't all agree in Computational Creativity!

It's not all Bad!

- In this afternoon's lecture...
 - Filling the humanity gap
 - Managing people's perception of creativity with the creativity tripod (now spider!)
 - Formalising progress in terms of the processes that software undertakes

Was there anything
right about your talk..?