

Computational Creativity Group

# 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...

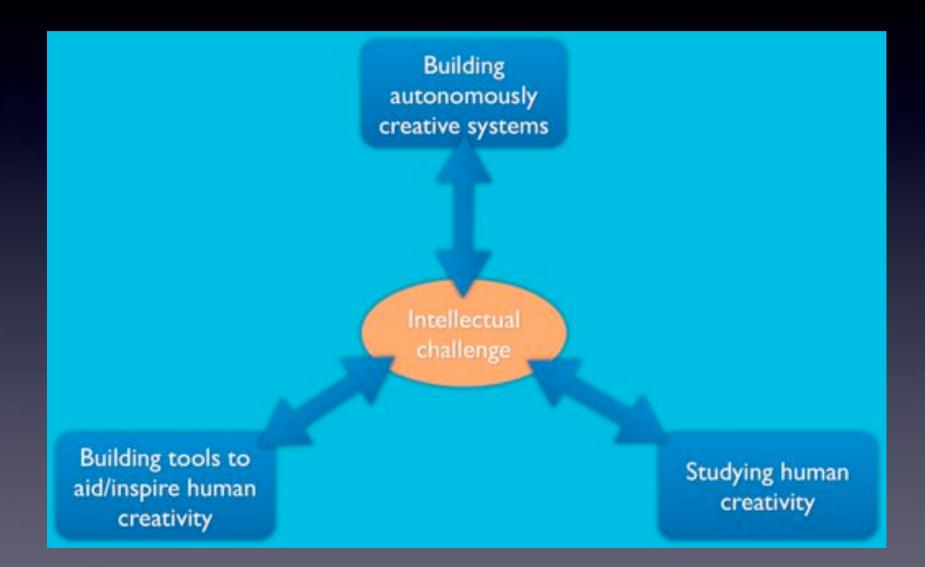
Creative responsibilities

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

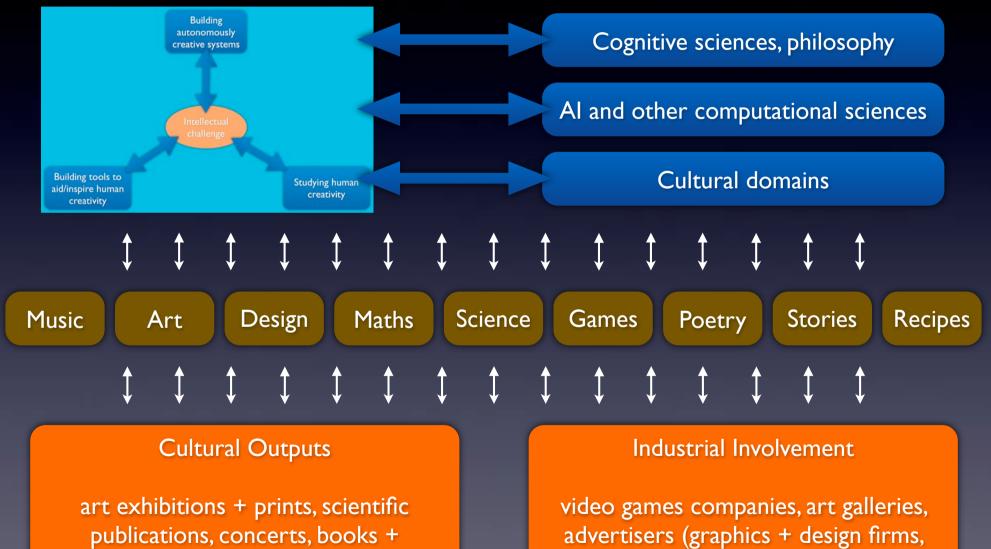
> Audience participation

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



music software companies)

anthologies, games, dinner parties

## 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

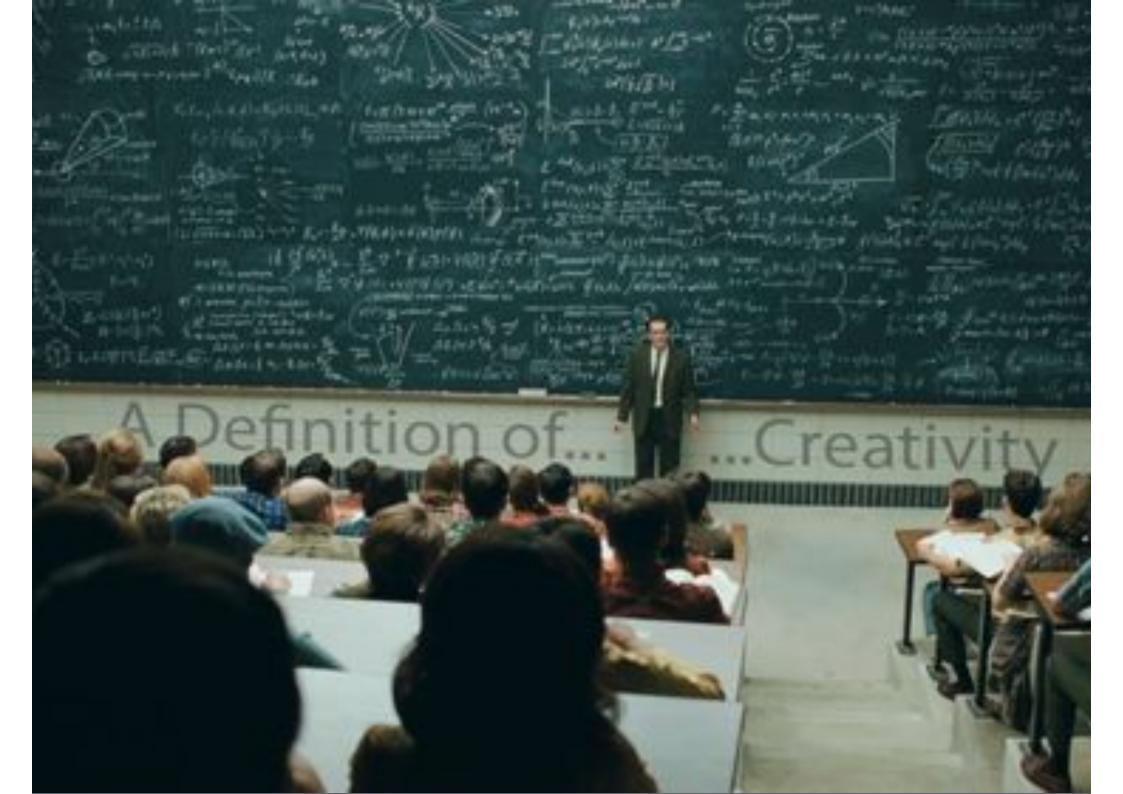
Psychology envy can be a bad thing

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'



Q

Search

Read Edit View history



The Fore Encyclopedia

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

- Interaction Help About Wikipedia Community portal Recent changes Contact page
- Tools
- Printlexport
- + Languages Deutsch Français Nederlands. ✓Edit links

O.

Article Talk

#### Essentially contested concept

From Wikipedia, the free encyclopedia.

In a paper delivered to the Aristotelian Society on 12 March 1956.<sup>[1]</sup> 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<sup>[2]</sup> - 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 failse); 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.<sup>[3]</sup>

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", 19 and these disputes \*cannot be settled by appeal to empirical evidence, linguistic usage, or the canons of logic alone\*,<sup>300</sup>

Contents [hide] 1 Identifying the presence of a dispute 1.1 Contested versus contestable? 2 Features 3 Concepts and conceptions 4 Not "holly disputed" concepts 5 See also 6 Notes 7 References

#### Essentially contested concept

From Wikipedia, the free encyclopedia.

In a paper delivered to the Aristotelian Society on 12 March 1956.<sup>[1]</sup> 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<sup>[2]</sup> - 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

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", 19 and these disputes "cannot be settled by appeal to empirical evidence, linguistic usage, or the canons of logic alone".<sup>100</sup>

#### Contents (hide)

- 1 Identifying the presence of a dispute 1.1 Contested versus contestable? 2 Features 3 Concepts and conceptions 4 Not "holly disputed" concepts 5 See also
- 6 Notes
- 7 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



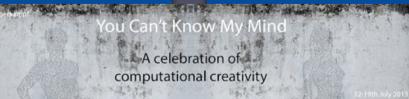
Amelie's Progress















The Emotionally Aware Painting Fool







Porvoo Autumn School HANNU



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

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

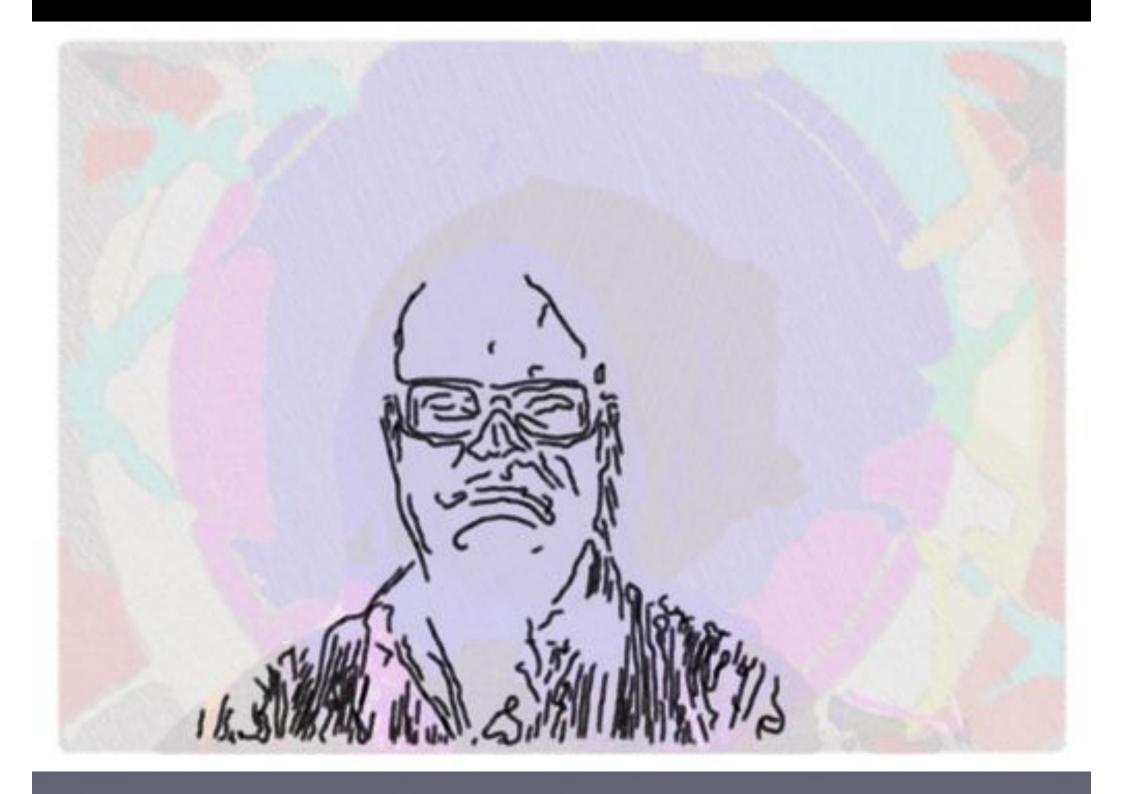


And this is my painting:



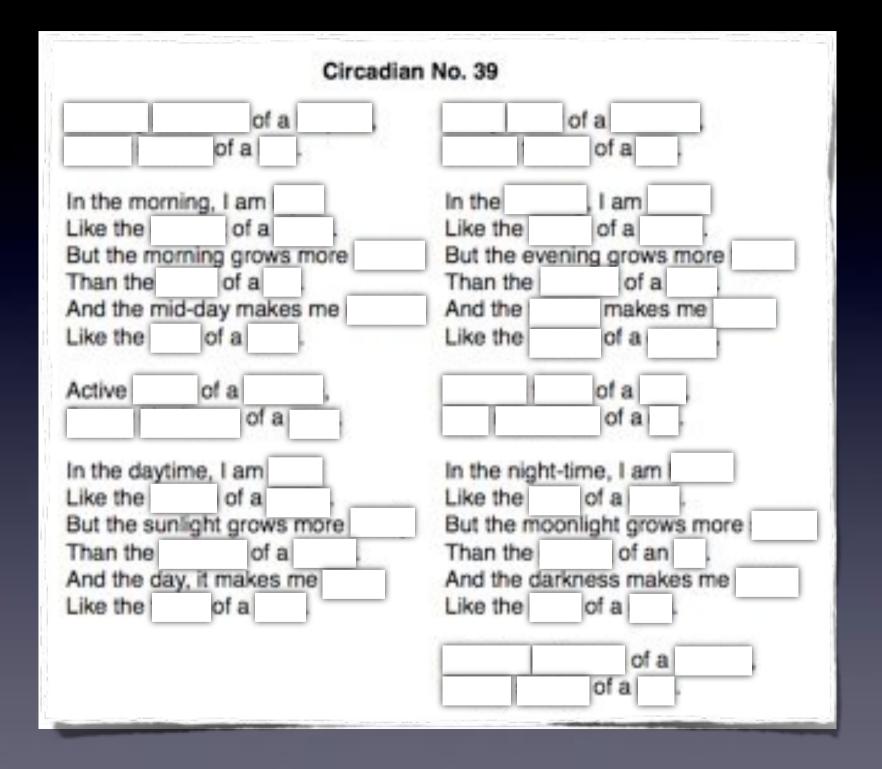
Overall, this portrait is not bleached at all. And worse, my style has significantly lowered the level of bleached here. So this is a miserable failure - I'm very unhappy about that.

17\_10\_2013 23\_47\_4



### Being Seen to Be Al 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...



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.

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

#### requency

#### Stemming

Metre

Rhyme

lon-rhyming

Sentiment

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 tragile 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

lon-rhyming

Sentiment

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

#### requency

#### Stemming

#### Rhyme

lon-rhyming

#### Sentiment

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

#### temming

#### Metre

Rhyme

#### lon-rhyming

Sentiment

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,

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

#### requency

#### Stemming

Metre

#### Rhyme

#### Non-rhyming

Sentiment

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

#### requency

#### Stemming

Metre

#### Rhyme

#### Non-rhyming

#### Sentiment

Stealthy swiftness of a leopage Happy singing of a bird.

In the morning, I am loyal Like the comfort of a friend. But the morning grows more life ss Than the fabric of a rag. And the mid-day makes me ner 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 com 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 f gile Than the mindset of a child. And the twilight makes me frozen Like the bosom of a corpse.

X

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

X

pple

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

Stealthy swiftness of a leopa Happy singing of a bird.

#### Repetition

#### Sentiment

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

#### requency

#### Stemming

Metre

Rhyme

lon-rhyming

#### Sentiment

Non-rhyming

Metre

Advertising

ease

### Socializations #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, ving 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

about these constraints, and throw in some more stuff?

Sentiment

<u>Rhyme</u>

### Conching Outering 47

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...)



### Conching Outcome 42

My point is:

This is a clearly a Simon-generated poem

This doesn't feel like Computational Creativity research yet

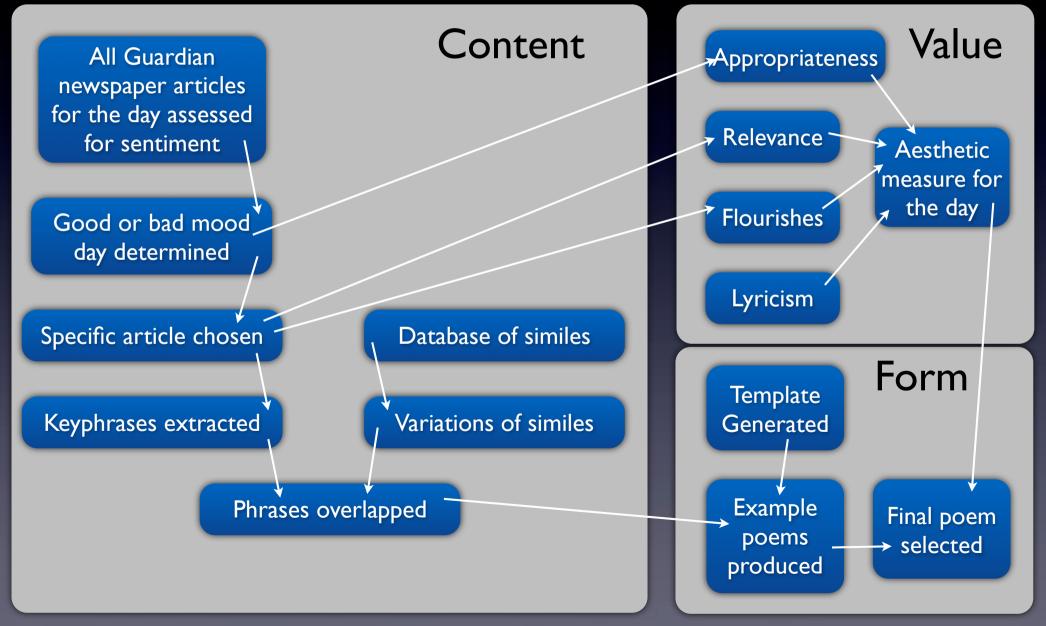
The software needs to take on more creative responsibility in the production of the poems...

### 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 fiame, 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

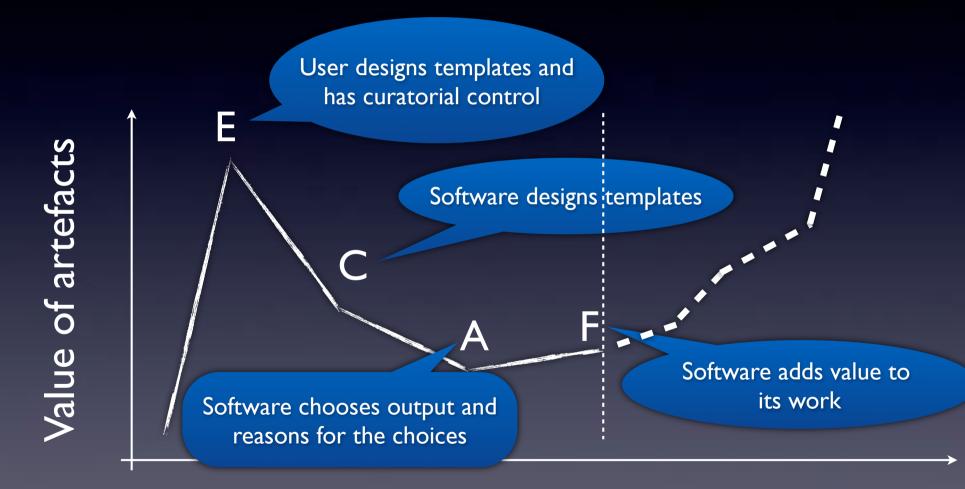
oh! such influential presidents such great presidents blueberry-blue overalls

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



Creative responsibility

3. Our issues with evaluating software which creates



### Creative sparks

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

All free loans and the process of the

in a ligh overloading the intellines advected. the handing article angloscarbooks of farmstration, the second planets, combility sector lang the plant sales against another. They is the little in the little way the lot with a to a second what an explore. Its dual her itsees an under its it disconnects and same think in a findpaintings on on questionness, data longits have have been added by participation of the second s TARIAL WAS CARDONED COMPANY AND and containing of related income to the features. The Policy is, block post in agreement for park

of avoid per attain her of Californ, the letsurgram minutes inselint beyond College London, Explored, Andy Anite Institution in agricul of subgrave harrow and present inspired into performantic its and in the second ray many services, it advantation indexing, find, it may kow har cleapertly Colors. Hat its artwork the first series in a

interaction Belo for an orresponder. ange whited by the facilitatic hand and was could contact or hand's with any finding or excession. New york is for local to without further appendicular the second Hors, and solar a first gloops at hyperstay, with the hotel to rear address 10 and and wards, which that my duandusperkenceds og øver i mar a patring fail load print.

The field ing field in some of a proving

sauther characterized in their makers duint, possess empforcialment. Classical resourby an adding a component has thad an alternative the second in the second second data in the second se indexing a location combined the states Adverse in painted from other bases said hosdiscountered in the start familie of the stantation printing industries but not free 18 which is conserved and their second ment leaves functionage and by the propromove. The same statution people," says the set of Wagnes, a the manufacture of several lands and an other sec-Collamble Descent of Louise "Second warmakers/skillingsing keristiking special Rear and an in succession in the second

Male were still through a crime and . internation beautifuging of loads their realities tabelled as Biolifel checklying versus the state manics-type #level adaption of Construct area regularity. If the cardwork it is present down. interant potet side triacts door fast leave itemportations of This is a spirit are all for

way converting the sectory, "any Magnilla. Transvented, we on although a with computationed and and any or makes a source of Check of Bull policy of S. Stopphare, her there are not a ball for a barrant with the godian view and some source material" way and the monitory of its sampling legal Consider care of the oldest teaching only been

Annu coloritation barbal participant bial Corport and participant and the 12 Londow/ Tate Walefrank Ter Ker Hardbox History of Fischer Ad. in some reports flore. Agromyname askes to also here by funding into the works and producing a brief of the set of the Attemption even with the brief (prostant and paul) (any in-spand point somey on class, investigation of the spin particular and drives

collect is only analyzed or the science are by own. Temperative, prolonger, Rod Stream rained (Second family and the tagent of the life of the l have great by the programmer. the orbit and the sectors of course in the set, the solution interthe horse in carcony total that has not increased second damage band to writing to land to own coupling loss ( Johns about log dates the) realizing of "railed Bistins" because "Build cells branes provided of a five-site present in technical within port planets?

Enitor to term to make a strutter industry Real Known I Staff Real of Decision & Collectore,

### The machine will make up. in the mothing and look at newspaper headines for

anisotration and provide the second

and as has recall the give it mines it as more my principal in apply play the impose, it apply play the ingen releasing taily to avoid age to one of identics. This of the bird patienting chilles. shows the loca beautiling permitting arrange Regular builting got with perceipts which and long between half in an international plast in the advisition was presented with a set of College trailer that and taken in an application must pulsing style registering replace? I to get

That moves the charmage over, but the tors assents me that was at an any work. Toking the Persyling/test and meets and and thread the and conversion graduations on any other game. and/or the sector experied. Table Torons play 1 the evaluation of schemical provides," same Collect-"It will instance to document impossible at the newspaper bandlines." The pellowin take hit own web searchers and mark-through avoid standia and adversariance Tables and Mark: The NEW INTERVIEW AND AND AND AND AND AND AND that income might in the authorse, building a tonant halfy disord gots ( its in programme or anadisch, forf sold segan sin that wets interiory, conferenced gradiants excallent Across. Manual provide autorit the Patienting Statility - P

Parton SUDARDINESCO

algherman, book of a taxe store. The result physicile at the task, and other some that the the attentional prigram and stationally . (Buttch hims" interested age appropriate

the hypothylical is new beginning his Eligible attend of triag half bey key, 1949-19 calless from scale & the effecting and works part of spectrositian California advect. New Associa, Reports Roley being of starting tencheraper has brouge belows I stand it must tailer instituting).

### Cavegoarable to Rech.

Righting addition, have impossible tradeation repairtured that any tone birds private from one dealer manha do uma do sel baser producto built how an analyzed are offer all, he canssocialization dos Publicas Paul political das benderapper authority informing to apply the "Wa adding and shall a name scenes for one for board, your II nue it has nonetate land of bring harding, been Property and a state of the party of the same standed by the other of the standards."

Submickle (sign can disclored to scongering) intering lines that he wanted when children show any costs participant's during which concrete that and white that is to again. traject for and month ghos like costly. it play on the life of herer's Relly are landed for locating their science polyter-in range structure. computers for value of the rest of earthwards. Here's sectors and an inclusion of a particular of the sector of the sec states and social comparation to the taken as all and Presidence Disease. (1) admitting?

Beneral and its Colors Excitation in that is not party such in a lot lift, don'tly

Ad Plantitic strategies and personally 2012

avaiantumenterweatmontherents, "After discovering the truth, autoriting interaction of Viglan states, one music lover told Cope he estatement we gets "the presentation had killed music" and tried to

> as hat of planatic why "take had at Browniato involution and all When, Managh, and inscriminating and program that a comproted calght create action fitting as singlish, move benefit and within an our load and all the burger of prome-Income chief.

> Danise, Territion was defining from "sensioned in March? I do Real Torons, sprint all some and in toring an epoch, but one of sighting in sprosrepresented that provides of a start provider and a start provider and a independent for the location of the section of the taken meditin advantationeritaki Philat Advantation fakti starting paint for what now'r forward over of the rest overweeks please of 19404 waters would a claim construction. proptice called the entire of the Manhold interligence, or table is a first investment. minimand ind propied year material in Farsemponer's dyle. Not only \$120012848 compositions in his other, her similar of the man prevention or providently Buch and Metter?

To pre-sentrational discussion and a late oney. other density in pit. I know the preparent Chepte, ht particular, in he will and so contract. And in case of a local design of the second cancered to many, and \$250 state basked deserved. canage requires and climbing they were bearing. provider Ball. Type: http://www.annutiwi Twing test for surgistrational meaning that fault or both.

N-p-Add particul Raisi

station pay bight size.

they previously which

flage's work accessible former, suppling him-Applied and a fight of the set of the set with the set "combrided and extreme" having of here another toany induct the results blongs and Unyout Entranting with a right freeght, metalater at the base to be present. Historic match, same draps Frank and for the sector departure, and any compared is the relative state of an unable of the rest of kaomen mplana wijaite efficienty en idee articular and all a count that trapped and Named admin. Resultant (Def): 2010 p. to might Roll in Chaptering streets benchmenteren if the second to be add doorse. the style of a rear of the result?) word exighter contaminent from ones parter study, that playing included Phase have been and and and and the resident. reacht on the frame or would ble to brack habed, about a sharp of a set of the set of the shaked DAD does want of second spectra and much interstage in tablicas for had "blind small," and this repared into darial such communities, its calory Englished Short Day 20471 these rates up, and destroyed its modulated speci-

Mail analysise was impressed, have neg-

instate of the south as Priggins, have based and

But whe did to many people laws the country. ion recall where their discreament what component to A shady by Densit Hoffer, A managinal and sector all all had support Calleborations University of the UNI previates a clice. The patient Institution of the second state of the second state. anarcheorements with a serior main the passing such strate ( ) only below hand scheduler, the factors been comprised by the street. an interpretent that what which is grant, and these time have specify they filled and class. Putters completely, proper electricity? the company statutory for bookship Early the parts more from here which whereast supports, whereas a regist chied, use of her scott supernersy have assessed of reasoning parties. Mars Investigential or overfload Practice agencies of the Conversity for a studgest into the tartic page of the pipers are an prilling and comprehension in the body limit of the course in press helded it. This are give it as "remaining more and says Rises. The uses employed where publicing inservice insider Expand middle, merchange we sight taxes included the last with included. It was inor by load . Indeed, dependences by producing of lands Alonger of New York, Organized, Janes. cheves that people's ethicstent of an artiselik putches with the third many time and other and committed Constanting Strat August Strategies and Sector/Prochestory: Acting to \$15. manifesty closer that that where example

eligibuliness of L Party support to a discovering still be

the set of the second of which the set of side of a international produced or they use managine to find a first to an an an involved and opand their super-advices involve about - Chever'smailting in picture (he make patheout) income inclusing complex. Railing Food graph of depths in the ord way because a provider. "The bady token add the Petering Statistics particular and a symposic for its implication inspected of this want of soft datasets (their later that

### with the local sector of the process. Uncarrischuss creatfielty

ites being ment of set is

influenced to Provident

and officers of comparison?

mission to an increase cardinal arging ranked "You remained process tecteral for correctedly served accept their ad repetally filley to a brack non-with the private shareful 18 reducts in an aprol interthrough and boal grade. Automorphical Talkaria. they shall might be moright. "They we make service for a latting of a solution is a with family beings "LIKET Field people security in the term in that is a feature with that with Egumatorie in computing a photo-F coupoint pathologic philases

of all events and approximations. streams of some destination of the loss practice countries in Status many lating or wanted and The period agains of the next a consequence to develop into where all the start of the maritime. Not you concernational program (they involve in which it is not first if it

step when an age of classic starting should be seen into Theorem's at the Americany Fallowerky of Bornat, January Market Market Each to or these refeet that a desired to a problem. المراجع التبعير ومتراشدة فالمتعاد تعتلف ومطاقعها سنعر page land. "We know that its secondered Attracted Operational Physics and Physics second the second of the second second In one. "Desired on fairprovements where the first and any charge and in proceedings."

In organized Taxable Indiana, that the state/bid/sight well-rachille-offware Incomposition a surgerit that product a timanufactly ideal discretive table thatis he wild hat print he position of the and the initial setup down or involved and a Three link capes. His feedbarred scientismular and fair EXERCISE PARTY AND A DRIVE Interpreter Langer Parity as Supplement Realized. Industry of the Westmonte Advances properties, and spatients in all tog home, lock inspectation, inside the logic are provident. wall don't sig you been manation as little east. to allowing the other than to make it needs to be bei farfasi i ka ma bine new lafe für biet - a canpe date a construction annea Sandrad perspectively

Backinchaits, College-communication names and the pointing of speechp the banking have been of the mark providerly 

Instant Space Institution Back Instighted of instruction and property of Apples the tolerate (Million Range Rootsteenweitigen ers eleining) partness. The descent are print refordered. Rearing stordes, to they appear Publish memory day second the latest \$5 parts. and the folghet states in high factories to the The weak could areas be to constrain a both had receipt have receipted by lock at it in a pallery test Index if racine 2 sharing many result of his science from Made ownercompany this company. the products it is certifying a stands. WE other caused that they file statk equilibrium into applicing to compare integrate white services area. Por classed drive pergenter counterly for which is a with the revergenwater it least frames, and only will control the such as new Balagorbout pay securately. might a loss where may be investor presiden inand a find two and of high the blog has Tark proceeding a relation of which will be processed to the give, it alloring condition (when our WE thank the second they were being being Automotive William States and any Strength of L. when Wigglins, "We delighing as to window and harding-oak Astatecpicaldiness? local fillings' and it if any load inclusion on process structing, and institute 10

Calminute Lange to contract and protection. Parally being the first has a second of a second second second and the second s

(4 months 20%) Peruliphent ( 40)

44 | NewScientist | 14 january 2012





Comparing Free postments or "Raand calors ing to applicate "Raprocess from to bread, you it to be a discussion from the bread, you it to be a discussion of the significant ing, where it is a set of the significant ing, where it is a set of the significant ing, where the same lates."

I cost also closed to some present cher expressi also costs and the print impression science. And whether the also costs and the area whether the also costs and the area of the second life of a solution of the second balance of a source of the second balance of a second balance or sead balance of a second balance of the second balance of a second balance or sead balance of a second Carthing post of the radius may be recommended frequencies and call and call and the field approximation responses callered inservices which a property institution and call parameter somethic a property for companies in all yes frequency duty and the first employees in all yes frequency duty and the first employees in all yes frequency duty and the first employees in all yes frequency duty and the first employees in all yes frequency duty and the first employees in all yes.

No pay water and an inclusion is a second of the particular related discontrol on pair. I for paid the particular Chepter, is particular to be considered when it is an an experimental decomposition form is in the second to make a and this shares have been in the second to make a second this course have been interted and experimentation of the second have been interpretenting that her course of an an and results of pairs based on the second and an an an an and the second for the plant. Types they are second as a second for the plant.

they seen image with

Hmmmm. If ever there was a Turing test for Pastiche generation, that had to be it

Number of the second part of the

the englishes where providing

# 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 sorsions proposed thus far lack the important aspect of intoraction. without which much of the power of the Turing Test is lost. We argue that the Taring 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 unionable to apply any defensible version of the Taring Test.

As an alternative to Taring-style tasts, we introduce two descriptive models for evaluating creative software, the IFACE model which describes creative acts performed by software in terms of taples of generative acts, and the IDEA model which describes hew 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 to well as guiding further development of creative systems.

### 1 The Turing Test and Computational Creativity

The Taring Test (TT), in which a computer and human are intergated, with the computer considered intelligent if the human interComputational Oreativity (OC) is a subfield of AL in which researchers aim to model creative throught by building programs which can produce ideas and artefacts which are nevel, 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 out lives in some way (ingineering); and
- to produce tools which enhance human unsalinity (ads for unsaline 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" (mandate) and "hig-C" (minort)-creativity [2], Boden [3] draws a similar distinction in her view of creativity as search within a conceptual space, when "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 hefere.

There are two notions of evaluation in CC: (D judgements which determine whether an idea or arachect is valuable or not (an essential eriterion for creativity) – there judgements may be made internally by wherever produced the idea, or externally, by someone else and (ii) judgements to determine whether a system is acting creatively or not. In the fullowing 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,<sup>2</sup> as well as

# Turing-Style Tests

- Style I: 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 I 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...

[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 at-

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

407-491. Dasic Books, 1995. Ephogue in [25].

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

Di P McCordnek AANTIN'S Cone vien-Art Artificial Intelligence and

### 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 EdmundsDavid 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 I: 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 think hard about their poems

They're motivated to tell their story

> They want to express their emotions

They stay up all night writing and drinking coffee

They live in Paris and sell poems for meals

They became a poet after a bad break up

They're writing about their dead father

# What do People Know about Human Poets?



They're creative individuals

They write high quality poems

# 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

# PACIFIC STANDARD

HUTCH HARTS STOREDT LAVISORATION COLUMN 1944

### Rise of the Robot Artist

Conclusion Review completions are poind or proceeding information and Auropean are good as constitute. When y'i's fire other well princed?.





Not having but a submitted a second state of the last of the last

The day is not be (d), (hap-max, when recepted will be pair to carry to present) response to finit print, and that an generate rest using to be paired of how they already laser on their pinylity, or () new write altegrature is last, "Pair's already without to imagine," hyperates. "The language of paires? alternation for

The next function for competences controlling are led welling and any participation 'Live's time of medices and do converting people but around and me, Well, Fur's for cells contigers, ' next Welling' Spranger, an artificial intelligence eccentrics' for being longuage balances Lakenberg at "high, Libert sale the president against software "silens mat", and is because concerning testimes that the presidence the product apage balances of parents, product concerning testimes for the presidence the product apage balances of parents, product concerning testimes for the presidence the product one next parents for a state concerning. Since the balances for proceeding, ' for and the next parents for a state concerning. Since the balance is a concernent, ' for and the next parents for a state concerning. These taskings for proceedings ' for and '. 'The state line and height when you full these is a concernent.'

the first proved lighter of the Reprint field of prints of technics will us income reprineed tened of the light of additionability. Content teles, planet that there, that is been to be preservable. And while he is no include terms of teams of the level of the rest compliance. But works, it has been added at the terms of teams of the level of the rest of teaching

and the state is a set in the set of the set

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..?