

CSCI 3725: Computational Creativity



Formalizing & Evaluating CC

Today's Outline

- Creativity as **Search**
- How Would You **Evaluate** a CC System?
- (A Few) **Classic** Ideas in CC Evaluation
- Four **PPPP**erspectives

Creativity as Search

Recall: Boden's Three Types of Creativity

New ideas are generated by...

- **Combinatorial**: combining familiar ideas
- **Exploratory**: exploring a space of concepts
- **Transformational**: transforming the search space

Recall: Boden's Three Types of Creativity

New ideas are generated by...

- **Combinatorial**: combining familiar ideas
- **Exploratory**: exploring a space of concepts
- **Transformational**: transforming the search space

*...all of these can be described as **search**!*

Wiggins' Interpretation

- Universe **U** contains all possible concepts

Wiggins' Interpretation

- Universe **U** contains all possible concepts
- Rules **R** define the acceptable conceptual space

Wiggins' Interpretation

- Universe **U** contains all possible concepts
- Rules **R** define the acceptable conceptual space
- Evaluation function **E** assigns a value to a concept

Wiggins' Interpretation

- Universe **U** contains all possible concepts
- Rules **R** define the acceptable conceptual space
- Evaluation function **E** assigns a value to a concept
- Method **T**_{R,E} for searching U with respect to R and E

Wiggins' Interpretation

- Universe **U** contains all possible concepts
 - E.g., all possible sequences of words
- Rules **R** define the acceptable conceptual space
 - E.g., those sequences that match a given meter
- Evaluation function **E** assigns a value to a concept
 - E.g., does the text express the desired emotion
- Method $T_{R,E}$ for searching U with respect to R and E
 - E.g., produce haikus using a generative grammar and expressions reflecting the desired emotion

How do we use this “Creative Systems Framework”?

- **U**, **R**, **E**, and **T** are all system-specific
- Mine, learn, or model:
 - **U** and/or **R** for acceptable cases from existing examples
 - **E** from recognized examples or from the user
 - **T** which use existing examples and their properties

Applying CSF to Boden's Theories

- **Exploratory**
 - A space defined by U, R and E explored by T
- **Combinatorial**
 - T can be based on recombinations of existing ideas
- **Transformational**
 - ?

Creativity as Search: Meta-level (Wiggins, Ritchie)

- A language L , in which R , E , T are expressed
 - $R, E, T \in L$

Creativity as Search: Meta-level (Wiggins, Ritchie)

- A language L , in which R , E , T are expressed
 - $R, E, T \in L$
- R and T can be interpreted and modified during runtime

Creativity as Search: Meta-level (Wiggins, Ritchie)

- A language L , in which R , E , T are expressed
 - $R, E, T \in L$
- R and T can be interpreted and modified during runtime
- Transformational creativity is thus exploratory search for a new R or T (or both)

How Would You Evaluate a CC System?

How would you evaluate a CC system?

Based on what we know about CC systems so far, **how would you evaluate a given CC system** (any domain)?

It would be easy to state a single word or phrase as your response, but that wouldn't be clear or objective enough if we want to put our ideas into practice...

The aim here is to clearly **state** one or more **metrics** that we can **quantify** and **compare**...

“Evaluation of CC Systems” Blackboard Discussion

1. What (quantifiable and comparable) **metrics** do you think would be useful to evaluate CC?
2. How would you reasonably **collect** this data? How might this data collection be **automated**?
3. Are some of your metrics **domain-specific**? Do they require other specific scenarios or conditions?
4. Do they assess different **types** of creativity? If not, how could we do so?

Add your thoughts into a single *Evaluation of CC Systems Blackboard* post (you can do this as a group, but make sure all of your names are included)!

 **A FEW**

Classic Ideas in CC Evaluation

Evaluating CC isn't Trivial

Is this a
sandwich?



Evaluating CC isn't Trivial



A foamy liquid with shrimp and meat slices served in a coconut with an ice cream scoop.



18 Criteria (Ritchie, 2007)

- Any creative system is based on some **existing examples**, in one way or another.
- **Inspiring set**: examples **known** to the program designer (or examples that drive the computation)
- Inspiring set can be used as a training set

(we saw this with the mere generation continuum)

18 Criteria (Ritchie, 2007)

Judge the artifacts produced by a system in terms of:

- **Typicality:** Is the artifact a recognizable example of the *target genre*?
- **Novelty:** How dissimilar is the artifact to existing examples (i.e., *inspiring set*) of its genre?
- **Quality:** i.e., value

18 Criteria - A Few Examples

What proportion of the...

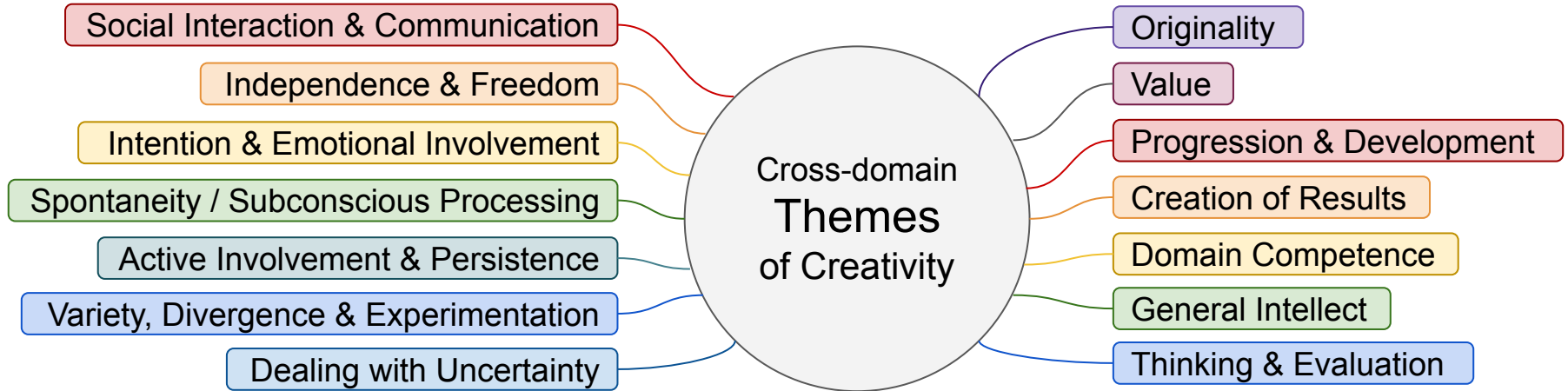
- ...**overall** results are **high quality**?
- ...**overall** results are **novel** and **typical**?
-**typical** (or non-typical) results are **high quality**?
- ...**novel** results are **typical and high quality** (i.e., demonstrate creativity within existing norms)?

Creative Tripod (Colton, 2008)

- One example of a claim that novelty, value, and typicality were **not enough**
- Creative systems must exhibit **skillful, appreciative** and **imaginative** behavior
- Colton claims **only** the audience's (**his?**) perception matters rather than how the system actually works (Jordanous, 2012)

SPECS (Jordanous, 2012)

Identified **14 themes** from 60 years of creativity research

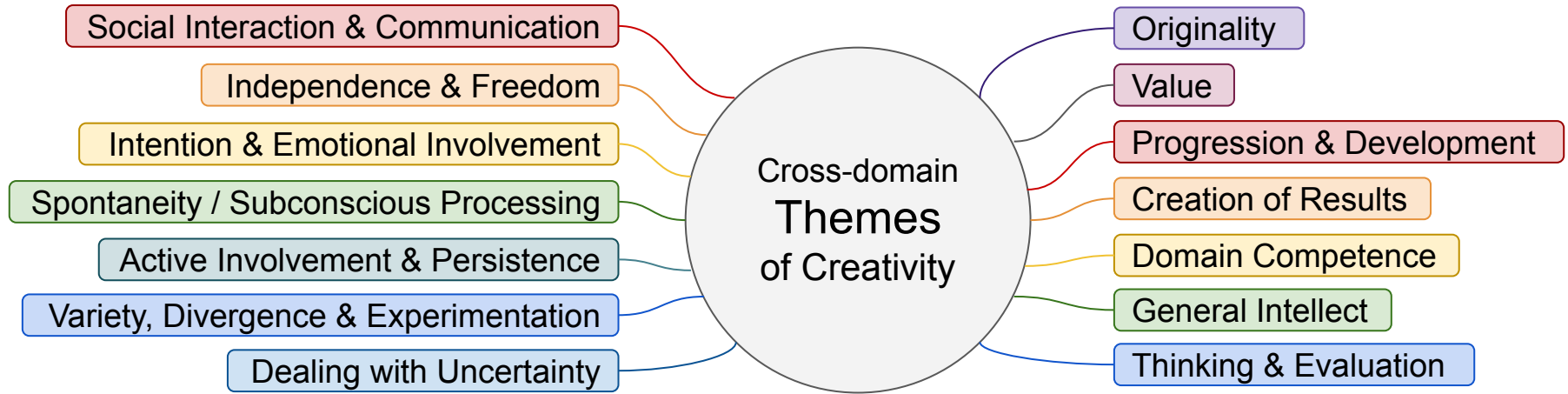


Breakout: In-SPEC-tions!

- Let's take a look at the **14 themes** together, and see if we can **distinguish** between them!
- As a group, consider the scenarios I give you.
- Determine what **theme** **best corresponds with each scenario**. (If you're not sure, you can state more than one with explanations.)

SPECS (Jordanous, 2012)

Jordanous decided on a **general procedure** for **evaluating creativity** based on prior work and these themes



SPECS (Jordanous, 2012)

Here's the general SPECS procedure:

- Identify a **suitable** CC **definition** that your system should satisfy (14 themes suggested as a base)
- Express **criteria** based on your definition
- **Test** your system against the criteria

Four PPP Perspectives

An Analysis of Creativity

The problems of modern life cry out for creative solutions.

Hence the growing interest in creativity.

But what is it?

Can it be taught?

Mr. Rhodes takes some of the fuzz off the
concept of creativity and assures us that it can
indeed be developed in children.

By MEL RHODES

1. Product

Product:

the **result** of the
creative process



2. Producer

Producer:

whoever **makes** the
Product

*(includes the computer
and those who use the
computer to create)*



3. Process

Process:

the **procedure** used by
the Person to develop
the Product



4. Press

Press:

the **context or environment** within which the Person operates to create the Product





Next Steps

- **Reflect** on your next steps for a better recipe generator with your team
- How might you **quantitatively assess** whether a new proposed iteration is better?
- How will your new recipe generator address the **Four P's** (Person, Product, Process, Press)?