

Vendredi 31 janvier, 2020

Que veut dire "comprendre" pour une machine ?

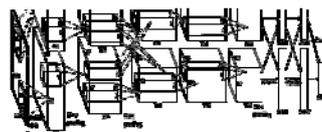
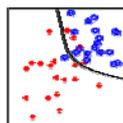
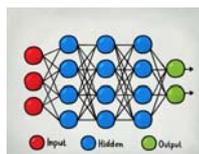


Jean-Louis Dessalles
Telecom Paris

jl@dessalles.fr
www.dessalles.fr



$A \& B \rightarrow C$



1970

1985

1995

2012

2019





What would "understand" mean for a machine?

- ❖ *A few challenges for AI*
- ❖ *What do AI techniques really do?*
- ❖ *Can we bring AI closer to natural intelligence ?*

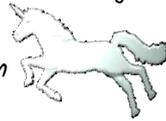
A few challenges for AI

Criminal investigation



One-shot learning

Unicorn



1,2,2,3,3,3,4,4,4,4



“prevent”, “around”, “together”, “abdicate”



<< Pain cuit toute la journée >>



“One swallow does not thirst quench”



“Delete all images that are duplicated”



Each context is unique

Criminal investigation



age of the victim
age of her children or parents
time of the crime
color of the room
...



→ decision

38% murdered women are killed by their partner

No understanding!

⚡ Limits of extrapolation from data

alibi ?

Each object is unique

One-shot learning



"buffet plate clip for wine glass"

*A child learns about four+
new words a day*

Goulden, R., Nation, P. & Read, J. (1990).
[How large can a receptive vocabulary be?](#)
Applied linguistics, 11 (4), 341-363.

 **background knowledge**

Lake, B., Salakhutdinov, R., Gross, J. & Tenenbaum, J. B. (2011). [One shot learning of simple visual concepts](#). *COGSCI-2011*, 2568-2573.

 **Statistical learning
can achieve
one-shot learning !**

Zhang, L., Xiang, T. & Gong, S. (2017).
[Learning a deep embedding model
for zero-shot learning](#). *ArXiv*, 1611.05088v3.

Lake, B., Salakhutdinov, R., Gross, J. & Tenenbaum, J. B. (2011). [One shot learning of simple visual concepts](#). *COGSCI-2011*, 2568-2573.

c)

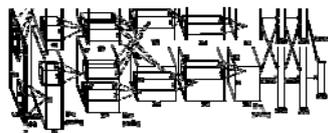
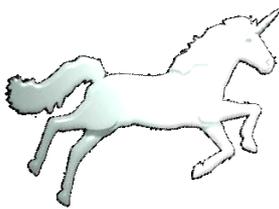
Stroke order:
1st 2nd 3rd 4th 5th 6th

Some level of understanding?

www.dessalles.fr

9

Each situation is unique



→ "horse"

No understanding!

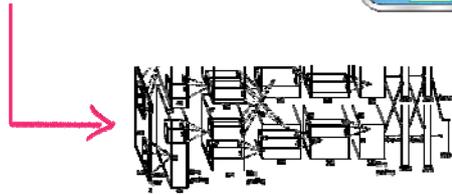
✱ Anomalies

www.dessalles.fr

10

Each pattern configuration is unique

1,2,2,3,3,3,4,4,4,4 . . .



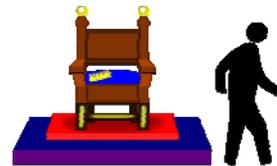
→ 1,2,2,3,3,3,4,4,4,4,4,4,3,4

No understanding!

✿ Structures

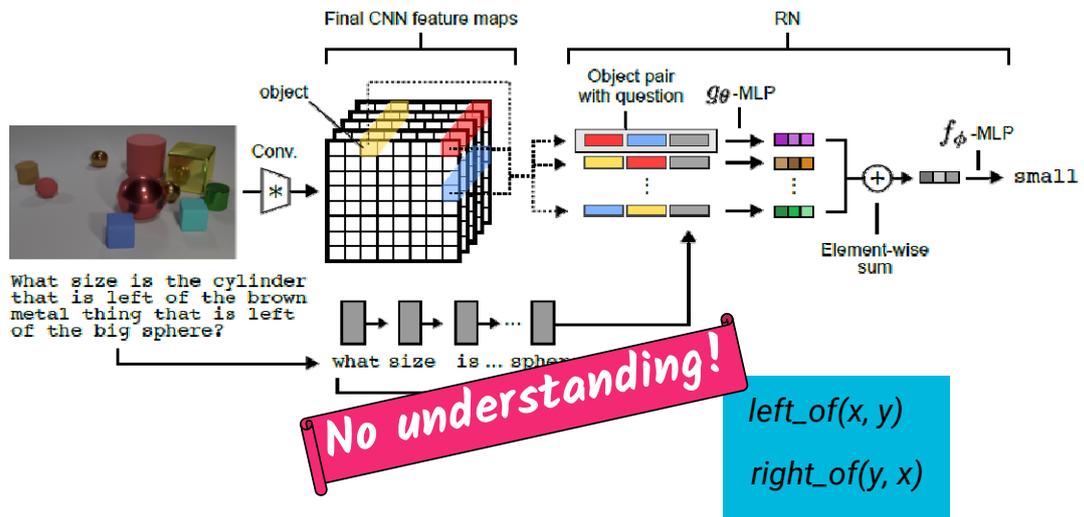
Invisible patterns

“prevent”, “around”, “together”, “abdicate”



✿ Relations

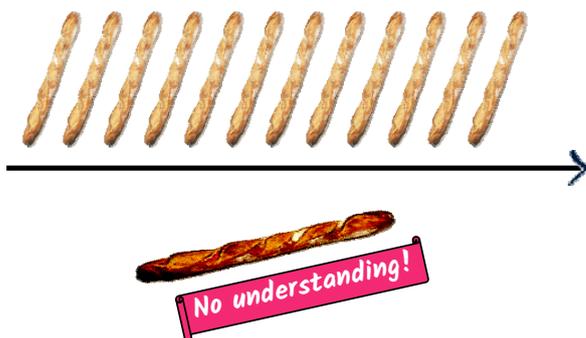
Santoro, A., Raposo, D. et al. (2017). [A simple neural network module for relational reasoning](#). NIPS 2017, 4967-4976.



www.dessalles.fr

13

Understand meaning



Victorri, B. (1998). *La construction dynamique du sens: un défi pour l'Intelligence Artificielle*. RFIA'98.



www.dessalles.fr

14

Understand meaning

“One swallow does not thirst quench”

(alluding to ‘One swallow does not a summer make’)

« Une hirondelle n’aspire pas la soif »



Hofstadter, D. R. (2018).
[The shallowness of Google Translate.](#)
The Atlantic, , Jan, 30.

❌ **semantic proximity ≠ semantics**

Understand meaning

In their house, everything comes in pairs. There's his car and her car,
his towels and her towels, and his library and hers.

Dans leur maison, tout vient en paires. Il y a sa voiture et sa voiture,
ses serviettes et ses serviettes, sa bibliothèque et les siennes.

Chez eux, ils ont tout en double. Il y a sa voiture à elle et sa voiture à lui, ses serviettes à elle
et ses serviettes à lui, sa bibliothèque à elle et sa bibliothèque à lui.

At home, they have everything in double. There is his own car and his own car, his own
towels and his own towels, his own library and his own library.

Hofstadter, D. R. (2018).
[The shallowness of Google Translate.](#)
The Atlantic, , Jan, 30.

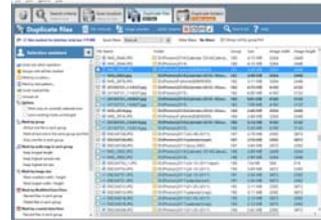
No understanding!

❌ **semantic proximity ≠ semantics**

Understand situations

“Delete all images that are duplicated”

No understanding!



relevance

A few challenges for AI

Criminal investigation



One-shot learning

Unicorn



1,2,2,3,3,3,4,4,4,4



“prevent”, “around”, “together”, “abdicate”



« Pain cuit toute la journée »



“One swallow does not thirst quench”



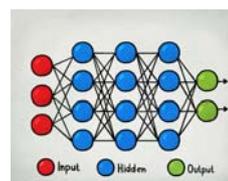
“Delete all images that are duplicated”



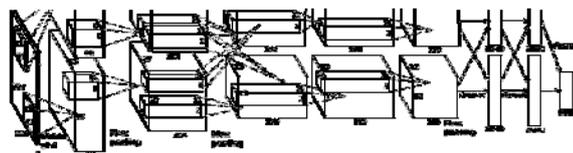
What would "understand" mean for a machine?

- ✿ *A few challenges for AI*
- ✿ *What do AI techniques really do?*
- ✿ *Can we bring AI closer to natural intelligence ?*

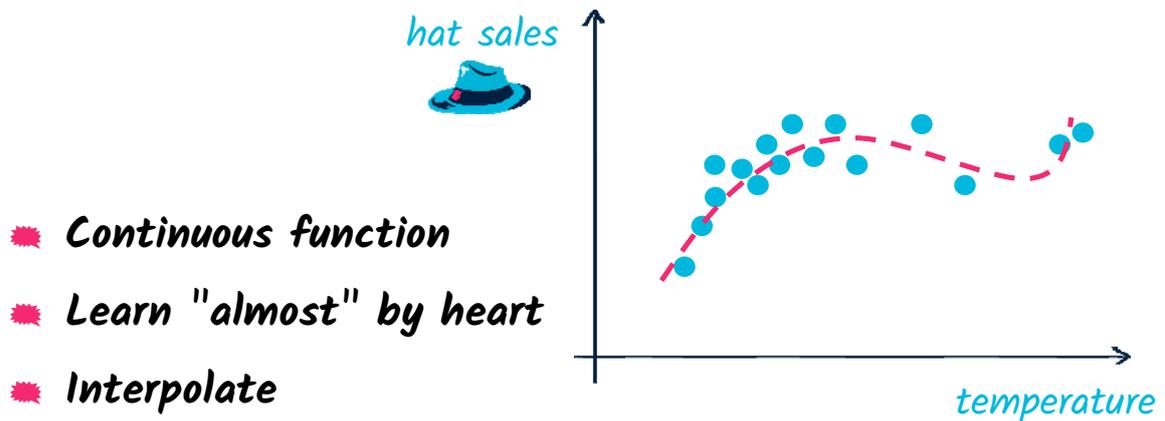
What do (current) AI techniques really do?



- ✿ *Continuous function*



What do (current) AI techniques really do?



Can we do better?



$A \& B \rightarrow C$



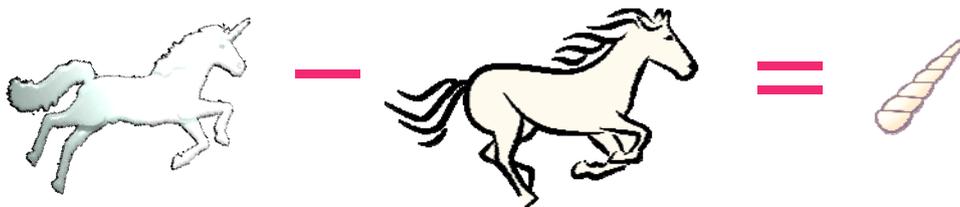
What would "understand" mean for a machine?

- ✿ *A few challenges for AI*
- ✿ *What do AI techniques really do?*
- ✿ *Can we bring AI closer to natural intelligence ?*

What would "understand" mean for a machine?

- ✿ *Can we bring AI closer to natural intelligence ?*
 - ⦿ *Contrast*
 - ⦿ *Simplicity principle*
 - ⦿ *... and other mechanisms*

Contrast

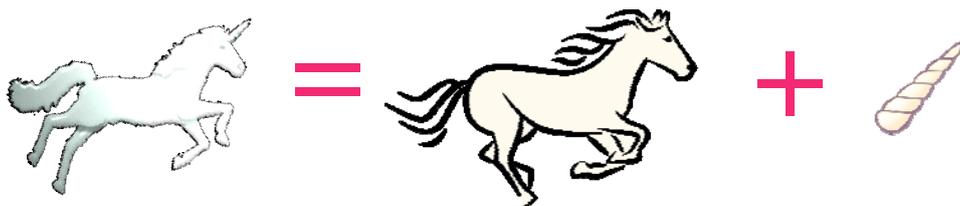


based on:

Gärdenfors, P. (2000).
Conceptual spaces: The geometry of thought.
MIT Press.

Gärdenfors, P. (2000).
Conceptual spaces: The geometry of thought.
MIT Press.

Contrast



based on:

Gärdenfors, P. (2000).
Conceptual spaces: The geometry of thought.
MIT Press.

Gärdenfors, P. (2000).
Conceptual spaces: The geometry of thought.
MIT Press.

Contrast

Dessalles, J.-L. (2015). [From conceptual spaces to predicates](#). In *Applications of conceptual spaces: The case for geometric knowledge representation*. Springer.

Predication: a small bacterium vs. a small galaxy

Negation: this isn't a sport

Systematicity: red apple, red pear

Metaphor: red face

Anomaly detection

Fiction: a flying pig

One-shot learning

Structure: AABABC

Analogy: queen : king :: woman : x

XAI: mortgage denial

Relevant descriptions: Elvis

Sileno, G., Bloch, I., Atif, J. & Dessalles, J.-L. (2017). [Similarity and contrast on conceptual spaces for pertinent description generation](#). *KI 2017*

Contrast

Dessalles, J.-L. (2015). [From conceptual spaces to predicates](#). In *Applications of conceptual spaces: The case for geometric knowledge representation*. Springer.

Predication: a small bacterium vs. a small galaxy

Negation: this isn't a sport

Systematicity: red apple, red pear

Metaphor: red face

Anomaly detection

Fiction: a flying pig

One-shot learning

Structure: AABABC

Analogy: queen : king :: woman : x

XAI: mortgage denial

Relevant descriptions: Elvis

Sileno, G., Bloch, I., Atif, J. & Dessalles, J.-L. (2017). [Similarity and contrast on conceptual spaces for pertinent description generation](#). *KI 2017*

Contrast



A red face is not really red, but the difference with a standard face is close to the prototype of redness.

contrast

redness



Contrast

Dessalles, J.-L. (2015). [From conceptual spaces to predicates](#). In *Applications of conceptual spaces: The case for geometric knowledge representation*. Springer.

Predication: a small bacterium vs. a small galaxy

Negation: this isn't a sport

Systematicity: red apple, red pear

Metaphor: red face

Anomaly detection

Fiction: a flying pig

One-shot learning

Structure: AABABC

Analogy: queen : king :: woman : x

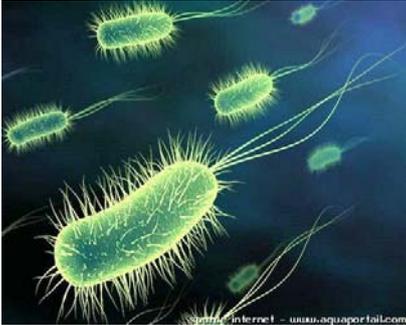
XAI: mortgage denial

Relevant descriptions: Elvis

Sileno, G., Bloch, I., Atif, J. & Dessalles, J.-L. (2017). [Similarity and contrast on conceptual spaces for pertinent description generation](#).

KI 2017

Predication



“a small bacterium” vs. “a small galaxy”

Contrast

Dessalles, J.-L. (2015). [From conceptual spaces to predicates](#). In *Applications of conceptual spaces: The case for geometric knowledge representation*. Springer.

Predication: a small bacterium vs. a small galaxy

Negation: this isn't a sport

Systematicity: red apple, red pear

Metaphor: red face

Anomaly detection

Fiction: a flying pig

One-shot learning

Structure: AABABC

Analogy: queen : king :: woman : x

XAI: mortgage denial

Relevant descriptions: Elvis

Sileno, G., Bloch, I., Atif, J. & Dessalles, J.-L. (2017). [Similarity and contrast on conceptual spaces for pertinent description generation](#). *KI 2017*

One-shot learning



“fuchsia”

Contrast

Dessalles, J.-L. (2015). [From conceptual spaces to predicates](#). In *Applications of conceptual spaces: The case for geometric knowledge representation*. Springer.

Predication: a small bacterium vs. a small galaxy

Negation: this isn't a sport

Systematicity: red apple, red pear

Metaphor: red face

Anomaly detection

Fiction: a flying pig

One-shot learning

Structure: AABABC

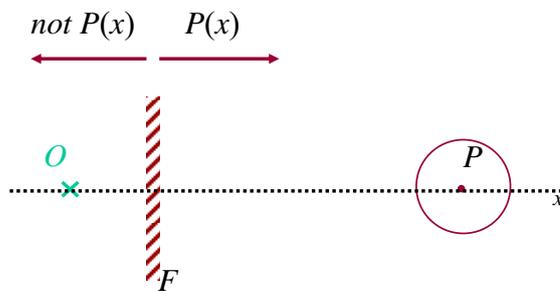
Analogy: queen : king :: woman : x

XAI: mortgage denial

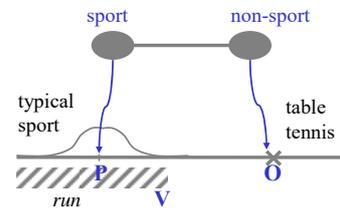
Relevant descriptions: Elvis

Sileno, G., Bloch, I., Atif, J. & Dessalles, J.-L. (2017). [Similarity and contrast on conceptual spaces for pertinent description generation](#). *KI 2017*

Negation through contrast: a dynamic process



[talking about table tennis]
 H- Oh ! That's not a sport !
 D- Well, you do sweat !
 H- If you don't have to run, it isn't a sport.



Contrast

Predication: a small bacterium vs. a small galaxy

Negation: this isn't a sport

Systematicity: red apple, red pear

Metaphor: red face

Anomaly detection

Fiction: a flying pig

One-shot learning

Structure: AABABC

Analogy: queen : king :: woman : x

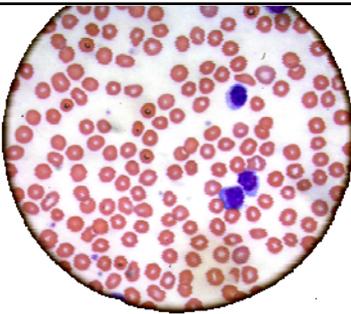
XAI: mortgage denial

Relevant descriptions: Elvis

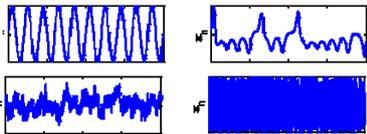
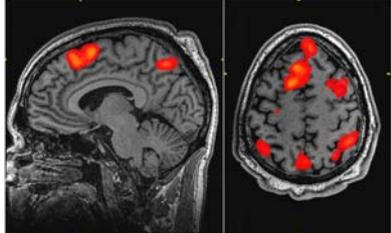
Dessalles, J.-L. (2015). [From conceptual spaces to predicates](#). In *Applications of conceptual spaces: The case for geometric knowledge representation*. Springer.

Sileno, G., Bloch, I., Atif, J. & Dessalles, J.-L. (2017). [Similarity and contrast on conceptual spaces for pertinent description generation](#).

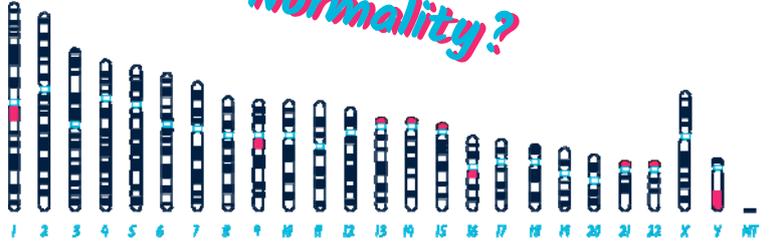
KI 2017



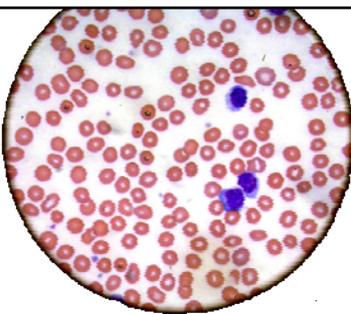
Anomaly detection


Normality?



37



Anomaly detection

Series of events $\{x_i\}$.

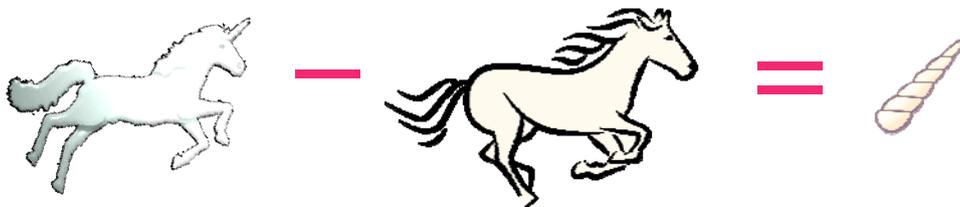
Contrast of x_n against $\{x_i\}$ \longrightarrow predicate f

Criterion: $\log(n) - C(f) + \log(|E|)$ $E = \{x_i | f(x_i)\}$

www.dessalles.fr

38

Contrast



*Contrast = a necessary step closer to “understanding”
but an ingredient is missing... Relevance!*

What would “understand” mean for a machine?

✳ *Can we bring AI closer to natural intelligence ?*

- ⦿ *Contrast*
- ⦿ *Simplicity principle*
- ⦿ *... and other mechanisms*

Simplicity principle

Chater, N. (1999).
[The search for simplicity:
 A fundamental cognitive principle?](#)
The Quarterly J. of Exp. Psychol., 52 (A), 273-302.



Complexity $C(s)$ of s :
 size of the smallest available
 description of s

$$C(x) = \min_p \{l(p) : M(p) = x\}$$

Kolmogorov complexity

"comprehension is compression"

Chaitin, G. J. (2004).
[On the intelligibility of the universe and the notions
 of simplicity, complexity and irreducibility.](#)
Grenzen und Grenzüberschreitungen, XIX, 517-534.

Simplicity principle

Cornuéjols, A. (1996).
[Analogie, principe d'économie et complexité algorithmique.](#)
Actes des 11èmes Journées Françaises de l'Apprentissage.

Murena, P.-A., Dessalles, J.-L. & Cornuéjols, A. (2017).
[A complexity based approach for solving Hofstadter's analogies.](#)
ICCB-WS 2017, 53-62.

abc → abd

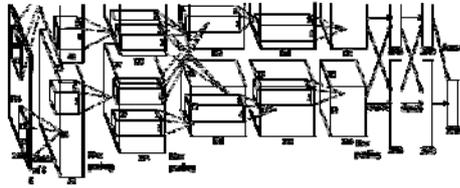
ppqrr → ppqqss

'ppqqss' = $\operatorname{argmin}_x C('abc', 'abd', 'ppqrr', x)$

(talk, talked) → (solve, solved)

✳ **Analogy making**

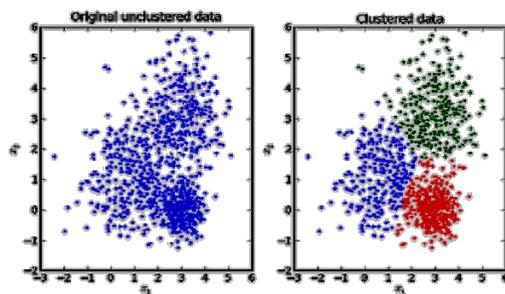
Simplicity principle



N objects, n classes

$\log_2(N)$ bits \rightarrow $\log_2(n)$ bits

(lossy compression)



www.dessalles.fr

45

Simplicity theory

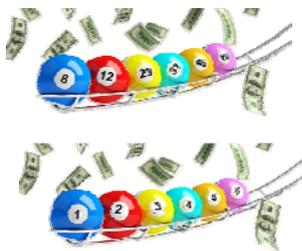
Unexpectedness = expected complexity – observed complexity

$$U = C_{exp} - C_{obs}$$

Dessalles, J.-L. (2006).

[A structural model of intuitive probability.](#)

7th Int. Conf. on Cognitive Modeling, 86-91.



Combinations	Complexity	Probability
1 2 3 4 5 6	3	$p/8 \times 10^{-6}$
34 35 36 37 38 39	6	$p/10^6$
10 11 12 44 45 46	11	$p/32760$
7 8 9 37 38 39	12	$p/16384$
8 9 26 27 28 29	12	$p/16384$
10 20 30 31 32 33	12	$p/16384$
1 2 5 6 15 49	14	$p/4096$
...
14 24 36 38 42 44	26	p

www.simplicitytheory.science

Unexpectedness

www.dessalles.fr

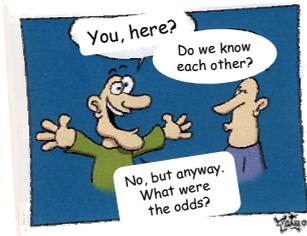
46

Simplicity theory

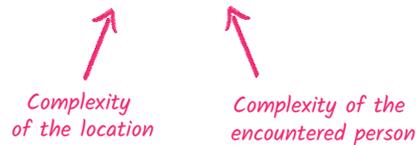
Unexpectedness = expected complexity – observed complexity

$$U = C_{exp} - C_{obs}$$

Dessalles, J.-L. (2008). [Coincidences and the encounter problem: A formal account.](#) *CogSci*.



$$U = C(L) - C(P)$$



Lincoln & Kennedy

Kern, K. & Brown, K. (2001). Using the list of creepy coincidences as an educational opportunity. *The history teacher*, 34 (4), 531-536.

"creepy coincidences" between Abraham Lincoln and John F. Kennedy



- "Lincoln was elected to Congress in 1846, Kennedy was elected to Congress in 1946."
- "Lincoln was elected president in 1860, Kennedy was elected president in 1960."
- "Both presidents have been shot in the head on a Friday in presence of their wives."
- "Both successors were named Johnson, born in 1808 and 1908."
- "Kennedy was shot in a car named Lincoln"



...

$$U(e_1 * e_2) = C_w(e_1 * e_2) - C(e_1 * e_2)$$

$$U(e_1 * e_2) \geq C_w(e_1) + C_w(e_2) - C(e_1) - C(e_2|e_1)$$

$$U(e_1 * e_2) = C(e_1) - C(e_2|e_1)$$

■ **To be explained:**

- ⊙ Role of close analogy
- ⊙ Role of mere association
- ⊙ Role of round numbers
- ⊙ Role of prominence

Simplicity theory

Unexpectedness = expected complexity – observed complexity

$$U = C_{exp} - C_{obs}$$

www.simplicitytheory.science

- **Rarity** $U \geq \log N - \log P - C(f) - C(r)$
- **Proximity** $U = 2 \times \log(R / d)$ $L = \operatorname{argmin}(C(L) + 2\log(d_L))$
- **Anomaly** $U \geq A(k) - C(f) - C(r)$ $U \geq C(H) - C(f) - C(r)$
- **Coincidences** $U = C(s_1) - C(s_2|s_1)$
- **Relevance** $C_w(f(s)) - C(f) > 0$
- **Responsibility** $C_w(s) - C_w(s \parallel a)$
- **Emotion intensity** $E = E_h + U$



What would "understand" mean for a machine?

■ **Can we bring AI closer to natural intelligence ?**

- ⊙ *Contrast*
- ⊙ *Simplicity principle*
- ⊙ *... and other mechanisms*

Cognitive mechanisms

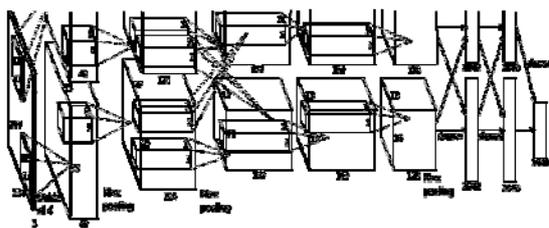
✿ Semantics

- ⊙ Contrast
- ⊙ Aspect
- ⊙ Semantic linking

✿ Relevance

- ⊙ Complexity drop (ST)
- ⊙ Conflict-Abduction-Negation

Precomputed Intelligence

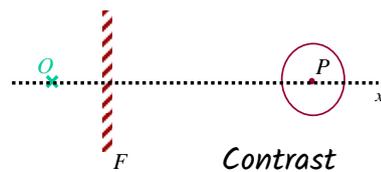


Pre-computed behavior

No understanding!

Intelligence "on the fly"

$\leftarrow \text{not } P(x) \quad P(x) \rightarrow$



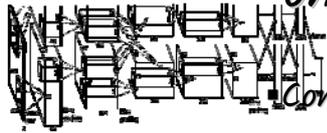
Simplicity Theory

$$U = C_{exp} - C_{obs}$$

Conflict-Abduction-Negation

Conclusion

Intelligence
"on the fly"



Contrast

A & B → C

▪ *Simplicity theory*

▪ *Conflict-abduction-negation*

▪ *Aspect*

▪ *and much more . . .*



Merçi pour votre attention



jean-louis@dessalles.fr

www.dessalles.fr

Visit: www.simplicitytheory.science

Special thanks to

- Jean-Bernard Auriol
- Laleh Ghadakpour
- Damien Munch
- Antoine Saillenfest
- Pierre-Alexandre Murena
- Giovanni Sileno