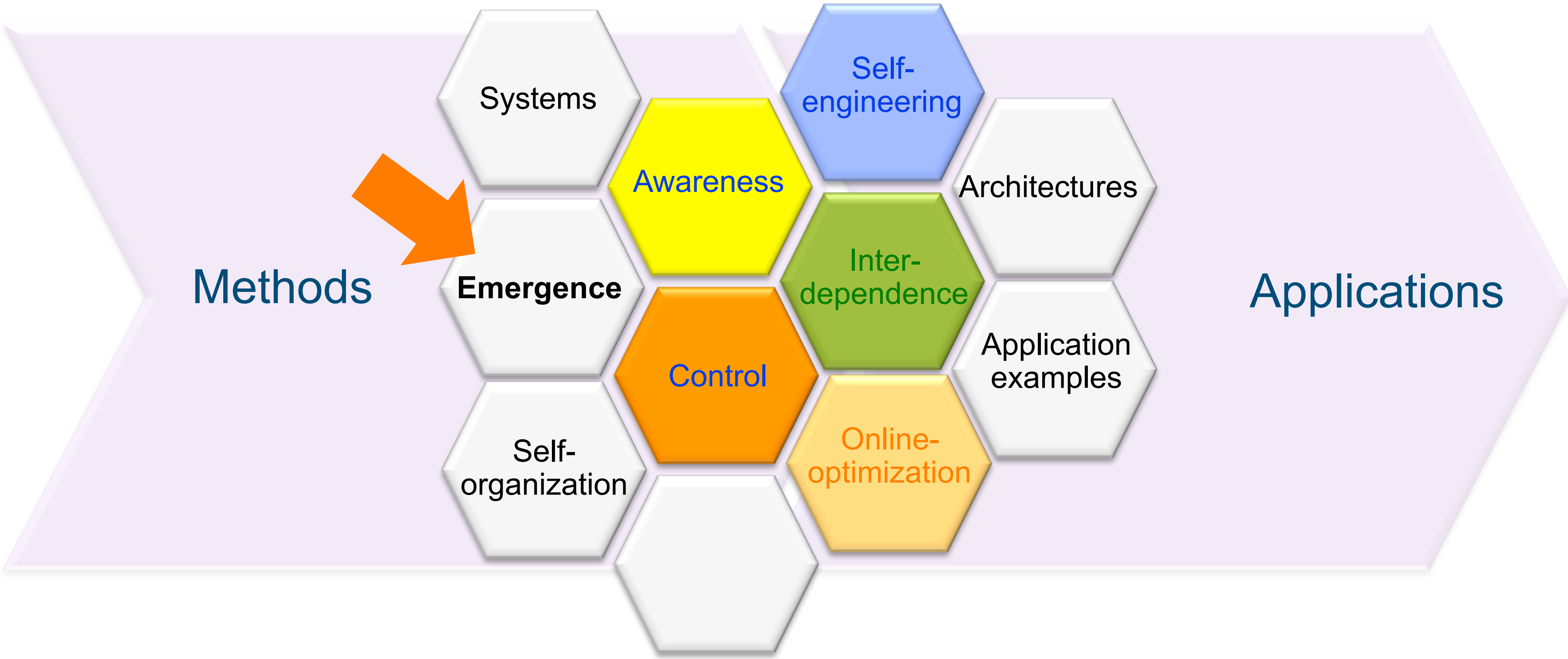

Methods
Systems
Emergence
Self-
organization
Awareness
Control
Self-
engineering
Inter-
dependence
Online-
optimization
Architectures
Application
examples
Applications

Preconditions

Information theory

Objectives

Derivation of a quantitative emergence measure

Content

- Examples
- Order and entropy
- Observation model
- Emergence definition
- Redundancy
- Emergence fingerprints
- Emergence prediction

Caution, this is work under discussion!

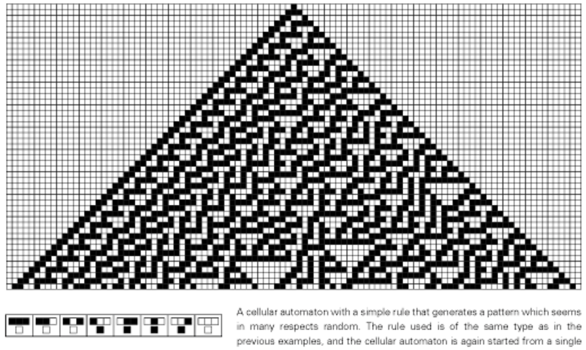
- Reading suggestions:

Mnif, Müller-Schloer Quantitative Emergence, Proceedings of the 2006 IEEE Mountain Workshop on Adaptive and Learning Systems (IEEE SMCals 2006), pp. 78-84, Logan July 2006

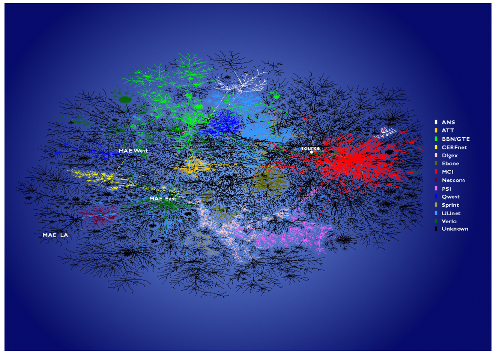
Fisch, D., Jänicke, M., Sick, B., Müller-Schloer, C. Quantitative Emergence – A Refined Approach Based on Divergence Measures; Proc. SASO 2010, Budapest, Sept. 2010; *Best paper award, pp. 94-103, ISBN 978-0-7695-4232-4*

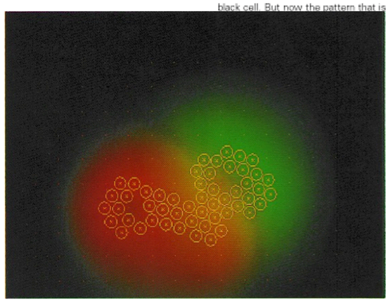
Examples of emergent structures

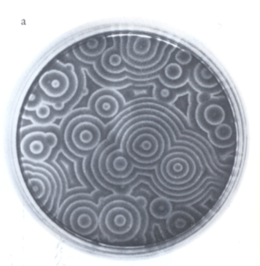
- ❑ Dissipative structures
- ❑ Ants
- ❑ Internet
- ❑ ...

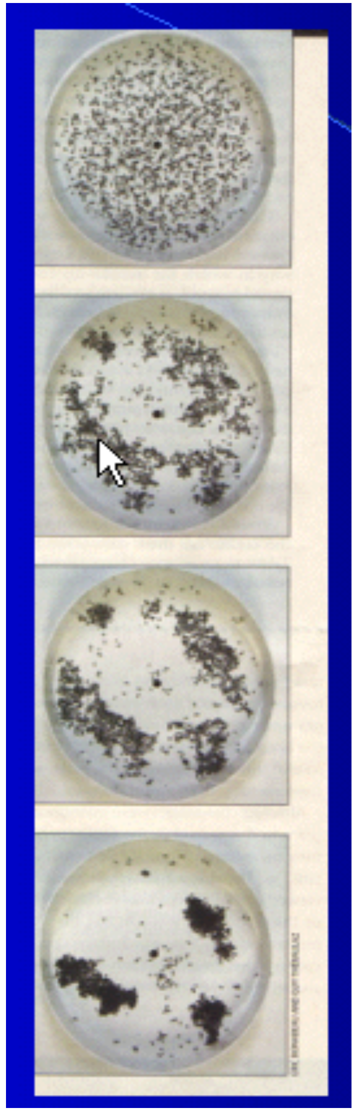


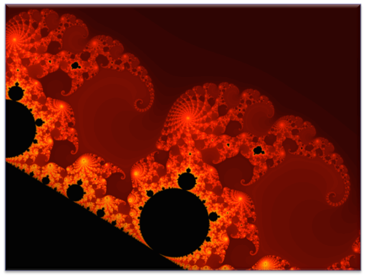
A cellular automaton with a simple rule that generates a pattern which seems in many respects random. The rule used is of the same type as in the previous examples, and the cellular automaton is again started from a single black cell. But now the pattern that is obtained is highly complex, and shows almost no overall regularity. This picture is our first
...enon that even with simple underlying rules and simple initial conditions, it is possible to
... In the numbering scheme of Chapter 3, the cellular automaton shown here is rule 30.



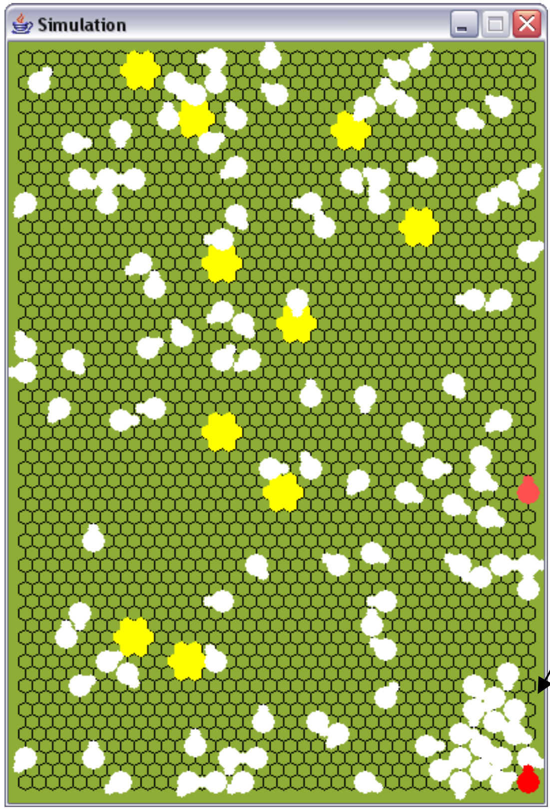




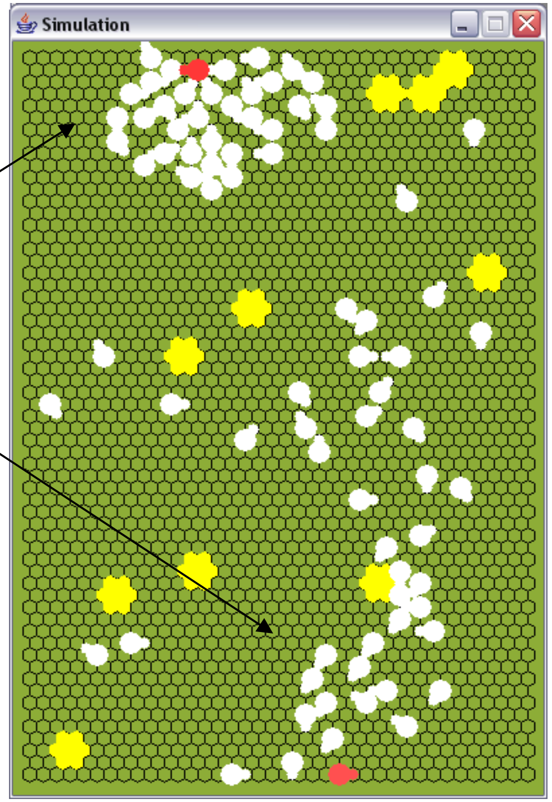




Another example of emergence: Chicken



„Undesired“ emergence



- Emergence from *to emerge (in German: „auftauchen“)*
- Characteristic of the whole system, not part of the subsystems.
- „A system is more than the sum of its parts.“ (Aristotle)
- Emergent system characteristics are not computable by *adding* the characteristics of the contained parts.
- Definition: An emergent system characteristic is a property, which is not only defined by the elements contained in the system, but by the interaction between these elements.

> - Example: Analysis of a living cell by cutting it into pieces destroys the life (destroys the process of interaction).
> - See also: Elan vital[1]

1) Élan vital: „Schöpferische Lebenskraft“ (Bergson 1907)

- Why should we measure emergence?
  - Emergence is neither good nor bad. But the presence of emergence indicates some source of order.
  - We want to recognise certain patterns to be able to react properly.
  - Beginning emergence can be an early indicator for some future behaviour. We can influence the system to support (positive) or impede (negative) emergence.
- How can we measure emergence?
  - Emergence measures the result (in terms of „order“) of some unknown process.
  - Order *per se* says nothing about self-organisation.
  - Entropy is a measure of order (or ignorance). How is emergence related to entropy?

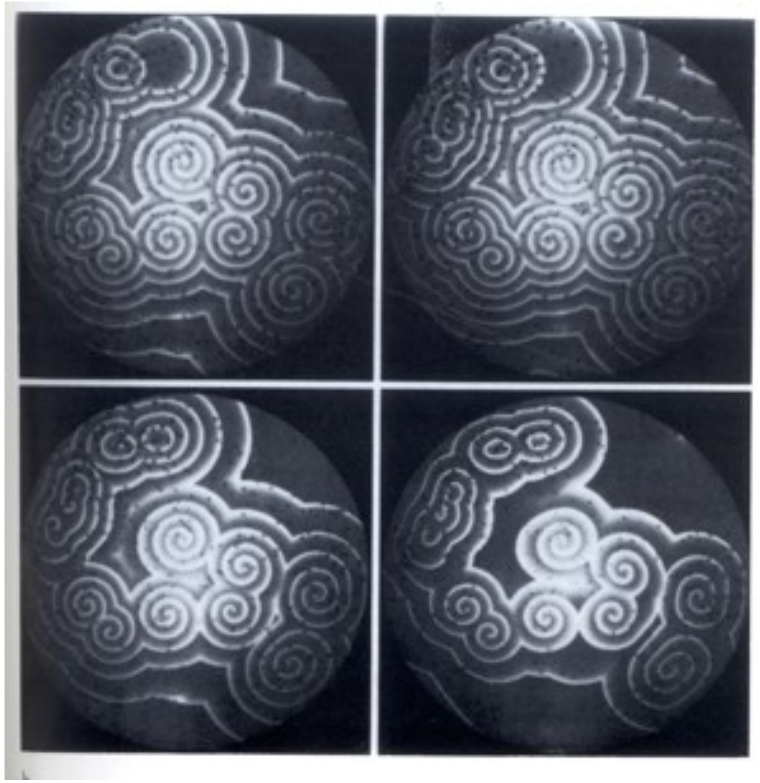
- Economies - beehives - financial markets - animal markings - team building - consciousness - locust swarms - mass hysteria - geese flocking - road networks and traffic jams - bacterial infection - town planning - evolution - the Web ...
- …these are all examples of emergent phenomena, where:
  - collections of individuals interact
  - without central control
  - to produce results, which are not explicitly "programmed”.
- ...and which are perceived as „orderly“.

Source: (http://beart.org.uk/?page_id=63/)

- Structural emergence in the sense of collective self-organization shows as:
  - patterns in time and/or space
  - patterns (order) at the system level, realized by interaction of (a large number of) similar individuals.
- These patterns have properties not existent in the individuals.

- Use verbal definitions of emergence as starting point.
- Goal: Assign a high emergence value to a system, which is perceived as emergent!
- Emergence is always associated with patterns (symmetry breaks).
  --> Structural emergence
- Patterns represent order.
- Order can be measured in terms of entropy.
- Therefore we must:
  - determine entropy and
  - relate it to „emergence".

What is order?

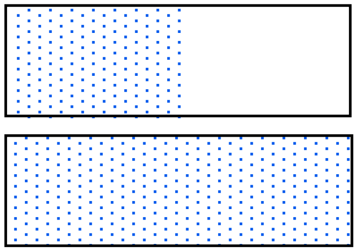


Dissipative structures (Prigogine)

Where is more order? Left or right?

Right: Higher entropy

Left: More structure



Gas molecules: distribution with low probability

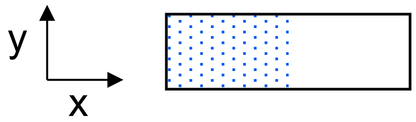
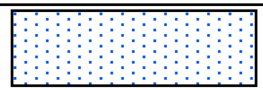
Gas molecules: thermodynamic equilibrium

higher entropy

Order is subjective

- The perception of order depends on the view of the observer.
- The purpose and the sensory equipment of the observer determine the view.
- A system can be rated as orderly or disorderly dependent on the utility.

- „Order“ or „disorder“ depends on
  - the purpose and
  - the view (aspect).
- A view (aspect) is determined by the selection of certain attributes (or a group of attributes) of an object.
- Example:

View =	x position	color
	higher order	same order
	lower order	same order

- The view is influenced by the preprocessing of sensory data.

Second law of thermodynamics: The entropy of any isolated system never decreases. Such systems spontaneously evolve towards thermodynamic equilibrium — the state of maximum entropy of the system.

- ❑ Entropy is a thermodynamic state variable.
  High entropy ⇔ high probability.
  - Clausius: Entropy of a closed system will always increase.
- ❑ Entropy is a measure of order/disorder (high entropy = low order).

$k_B = 1{,}38 \cdot 10^{23}\,\mathrm{J/K}$ (Boltzmann constant)

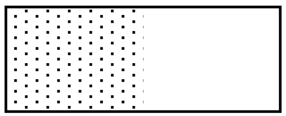
$$\boxed{\text{Entropy } S = k_B \cdot \ln(\Omega)}$$

- ❑ Statistical definition of entropy:

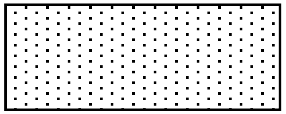
k_B = average kinetic energy of an ideal gas particle at a temperature of 1 Kelvin

Ω = probability of a macroscopic state (= the number of possible states of the particles in a system / total number of possible states)

- ❑ Example: irreversible diffusion



Lower Ω, lower entropy



Higher Ω, higher entropy

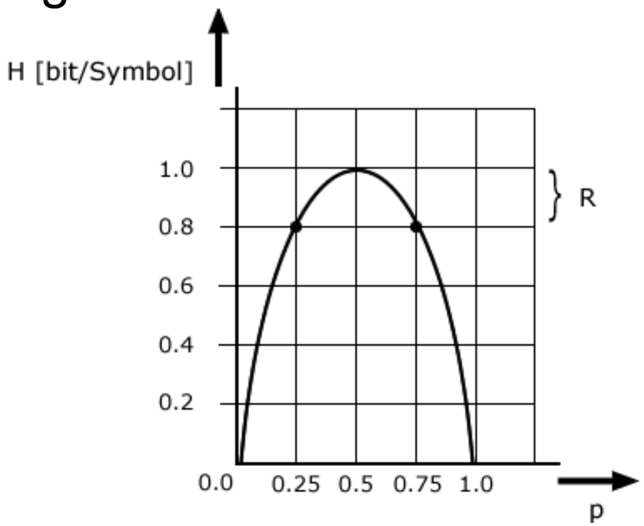
- Entropy is a measure of information (information theory, Shannon).
- Message source M, alphabet Z:

$$Entropy\ H(M) = -\sum_{j=1}^{|Z|} p_j \cdot ld\ p_j$$

- p_j = probability for occurrence of symbol z_j from Z in M
- H is a measure for uncertainty in a system (or a message source M).
- High content of random information ⇒ high uncertainty ⇒ low predictability ⇒ low probability ⇒ high information content.

- Example: 2 symbols (0, 1) with probabilities p_0 and p_1

  $H = - p_0\ ld\ p_0 - p_1\ ld\ p_1$



I thought of calling it "information", but the word was overly used, so I decided to call it "uncertainty". [...]
Von Neumann told me, "You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage.“

Conversation between Claude Shannon and John von Neumann regarding what name to give to the attenuation in phone-line signals.

- $H = H_{max}$ is desirable, if a channel must transport the maximal „newness" value per (physical) step.

- If $H < H_{max}$: $R = H_{max} - H$: The channel transports useless information (redundancy R).
  - Known information burdens the channel but does not increase the knowledge of the receiver.

- A Shannon channel is "good", if it transports the maximum amount of information

  $\rightarrow R = 0$

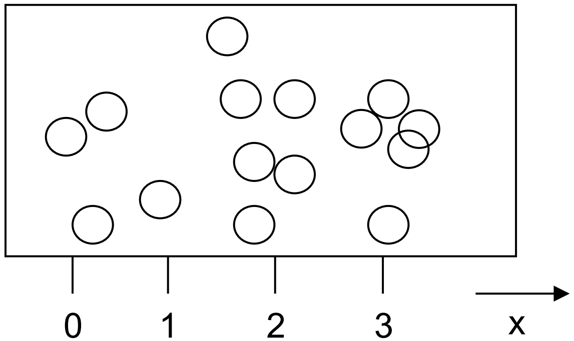
- Use the statistical definition of entropy
- Procedure
  - Select an attribute A of the system (discrete, enumerable) with values a_j.
  - Observe all elements e_i and assign a value a_i to each e_i.
  - Transform into a probability distribution over the attribute values a_j.
  - Determine the entropy $H_A = -\sum_j p_j \operatorname{ld} p_j$

- Each attribute X has its entropy H_X.
- System entropy $H_S = \sum_X H_X$

- Characterization of a system by:
  a) System entropy (low specificity)
  b) Vector of attribute entropies („fingerprint"): (H_A, H_B, H_C...)

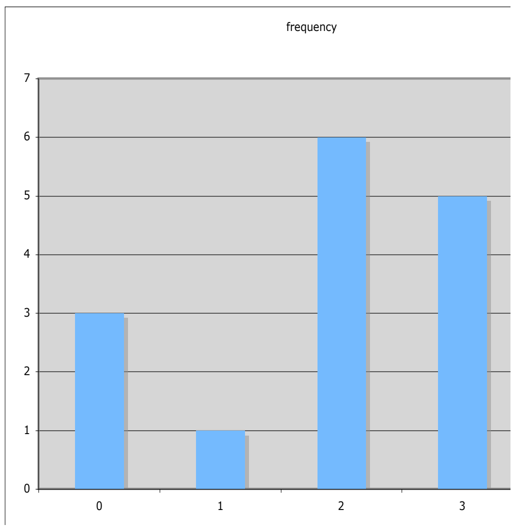
Example: Entropy of attribute „x coordinate“

- Discrete values of x coordinate: 0, 1, 2, 3



Position:	0	1	2	3
Frequency:	3	1	6	5
p	3/15	1/15	6/15	3/15

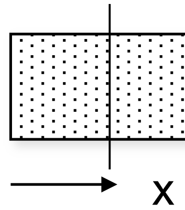
N = 15



$$H_{x\,coordinate} = -\left(\frac{3}{15}\,ld\frac{3}{15} + \frac{1}{15}\,ld\frac{1}{15} + \frac{6}{15}\,ld\frac{6}{15} + \frac{5}{15}\,ld\frac{5}{15}\right)$$

$$= 1{,}72\ bit/element$$

- System entropy and the entropy fingerprint are observer-dependent.
- Definition: An observation model determines:
  - which attributes are observed and included into an observation (selection),
  - with which resolution (i.e. abstraction level or quantization).
- The observer model is subjective and purpose-dependent.
- Example

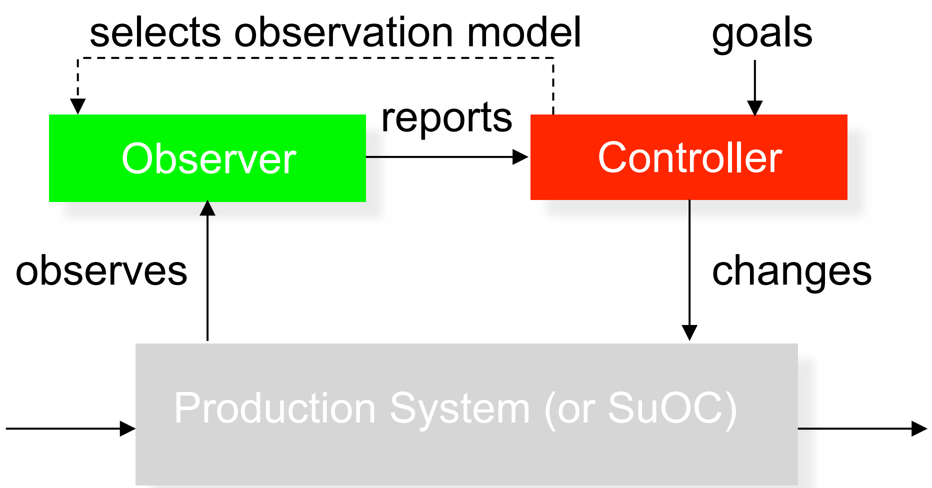


Attribute = x coordinate

  - Observation model 1: x measured as floating point variable (single precision): 32 bit
  - Observation model 2: x measured as left (0) / right (1): 1 bit

Entropy measures the information content of a system description.

- Entropy depends on the preprocessing of the observed raw data.
- Who determines the observation model? The controller!



- Classification of observed attributes
  - Basic: location, state, color ...
  - Derived (preprocessing): velocity, acceleration, average values ...

- Entropy ≠ Emergence!
- First try of a definition: Emergence M is the decrease of entropy H from a start state to an end state:

  $M = \Delta H = H_{Start} - H_{End}$
- This is a necessary, not a sufficient condition for a process to be called emergent. Emergence manifests itself by an increase of order,

  i.e. $H_{End} < H_{Start}$ and $\Delta H > 0$.

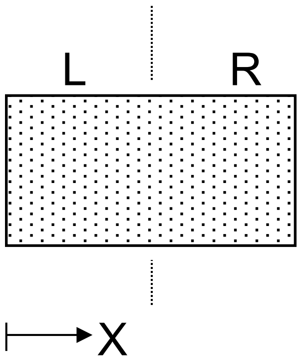
  In addition, the process, that leads to the increase of order, must be self-organised (not e.g. human-induced).
- Problem: The observation of emergent phenomena frequently involves a change of abstraction level.

Emergence Definition 2

- A change of view to a higher abstraction level leads to a positive ΔH, which is **not** due to an emergent process.
- $\Delta H = \Delta H_{view} + \Delta H_{emergence}$

  $Emergence\ M = \Delta H_{emergence} = \Delta H - \Delta H_{view}$
- If the two observations (start, end) are made on different abstraction levels, the increase of order due to the change of view (ΔH_{view}) must not be counted as entropy stemming from emergence.
- Or: The entropies H_{Start} and H_{End} are only comparable if they are observed on the same abstraction level ($\Delta H_{view} = 0$).

- Example



Observation 1 of x coordinate: 32-bit floating point

Observation 2 of x coordinate: quantization to 256 values (8 bit integer)

$$\Delta H = 24 \;\; bit / Element$$

$$\Delta H = \Delta H_{view}$$

$$Emergence\, M = \Delta H_{emergence} = 0$$

- Emergence $M = \Delta H_{emergence}$ is a measure of the (self-organized) increase of order from start to end of a process. Is there an absolute indicator of order?
- We can define order as the difference between the entropy at maximum disorder (H_{max}) and at a certain system state (H):

  $M = \Delta H = H_{max} - H = \Delta H_{emergence} - \Delta H_{view}$

  with $\Delta H_{view} = 0$:

  Emergence $M = \Delta H_{emergence} = H_{max} - H$
- H_{max} is the system or attribute entropy for the case of a uniform probability distribution.

- The Kullback-Leibler divergence (short KL divergence), Kullback-Leibler entropy, Kullback-Leibler information or Kullback-Leibler distance
  - according to Solomon Kullback and Richard Leibler
- define a measure for the difference of two probability distributions P(x) and Q(x) (or p(x) und q(x)):

$$D(P\|Q) = KL(P,Q) = \sum_{x \in X} P(x) \cdot \log \frac{P(x)}{Q(x)}$$

$$D(P\|Q) = \int_{-\infty}^{\infty} p(x) \cdot \log \frac{p(x)}{q(x)}\, \mathrm{d}x$$

- Emergence is the increase of order due to self-organized processes between the elements of a system.
- Higher order is measured in terms of a lower description complexity, i.e. in terms of lower information content.
- The more structure a system displays, the less explicit information is necessary to describe it: Kolmogoroff complexity.
- Any entropy decrease due to a change to a higher abstraction level (ΔH_{view}) must be subtracted.

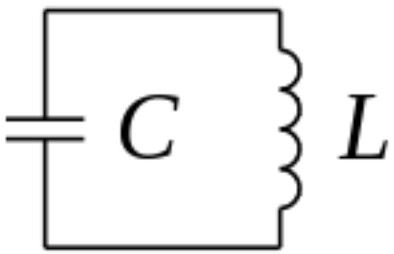
- Emergence $M = \Delta H_{emergence} = H_{max} - H - \Delta H_{view}$

- Relative Emergence m (for $\Delta H_{view} = 0$) $m = \frac{H_{max} - H}{H_{max}}$

- A quantitative determination of emergence is only possible
  - if the entropies (start, end) can be determined on the same abstraction level (under the same observation model)
  - or if the change of abstraction level can be quantified.

- If an emergent process results in totally new phenomena to be observed only on a high abstraction level, we can still calculate a ΔH but it might be impossible to separate this ΔH into $\Delta H_{emergence}$ and ΔH_{view}.

  Example: resonance frequency



$$f_0 = \frac{1}{2\pi\sqrt{LC}}$$

- Information theory:

  Redundancy $R = H_{max} - H$

  rel. redundancy $r = \frac{H_{max} - H}{H_{max}}$

- Why is $M = R$ (or $\Delta H_{emergence} = R$)?

- Answer: Different notions of utility
  - Information theory
  - Self-organizing systems

Redundancy (2)

- Communication engineer: Redundant information should not be transmitted over a channel. Redundancy is „bad“ (to be avoided).
  - A channel should be used only for the transmission of relevant (= new) information. Max. information content if $H = H_{max} \rightarrow R = 0$

    Newness or unpredictability of a message increase their value (information content). The physical steps (light pulses, current levels, voltage levels) should be used economically only to transport relevant information.

- Living system: The existence of predictable information about a system means that no (or little) new information has to be sent via the channel. High predictability (= emergence = redundancy) is „good“.
  - (Self-organizing) systems are „better“ if they display higher order, lower newness, higher predictability. This is the perspective of an “animal”.

Examples (1)

- Temperature is not an emergent property of a system. It is a measure of average kinetic energy of particles and involves a change of abstraction levels, no self-organization.

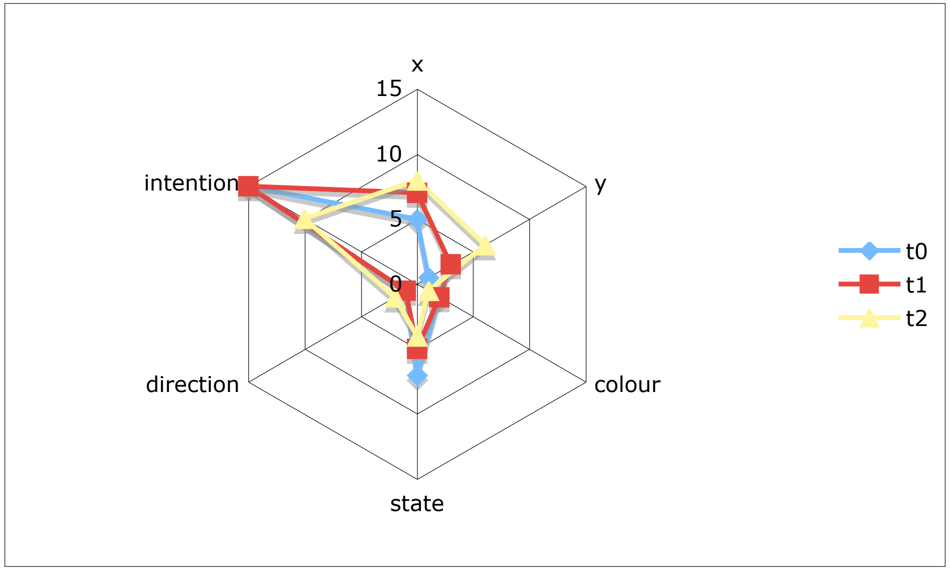
Examples (2)

- Balls in bowl
  - 70 black, 20 white, 10 red
  - $p_s = 0.7$, $p_w = 0.2$, $p_r = 0.1$
  - Observation model: color
  - $H_{max} = \text{ld } 3 = 1.58$
  - $H = 1.14$
  - $\Delta H_{view} = 0$
  - $M = \Delta H_{emergence} = 1.58 - 1.14 = 0.44$ bit/element
  - $m = \Delta h_{emergence} = 0.44/1.58 = 28\ \%$
  - But: This could be called emergence only if the color change was due to a self-organizing process!

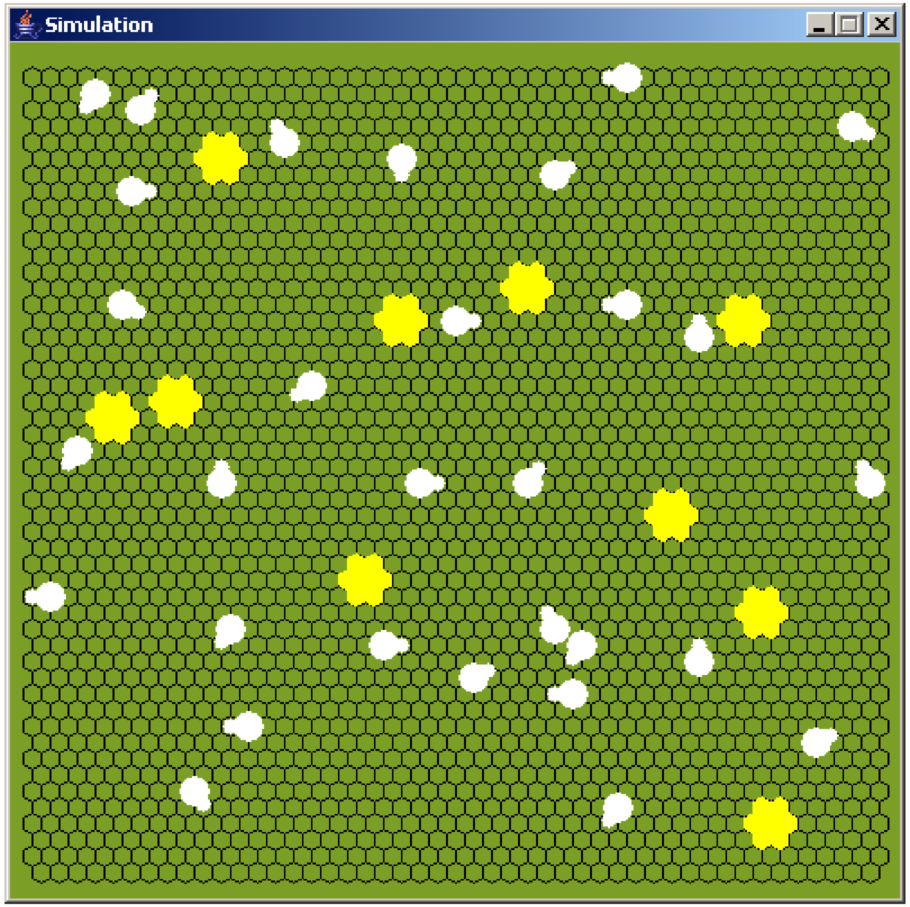
- Gas molecules
  - **View**: N molecules, only local coordinate x is of interest. x is (somewhat extremely) coded as 0 if a molecule is in the left half of the volume and 1 if it is in the right half.
  - **State 1**: Uniform distribution (thermal equilibrium).
    Entropy $H = H_{max} = 1$ bit/element.
  - **State 2**: All molecules are in the left half. Entropy $H = 0$.
  - Emergence ($1 \rightarrow 2$): $\Delta H_{emergence} = 1$ bit/element.
  - This can be called emergence only if the deviation from the uniform distribution was due to a self-organizing process!

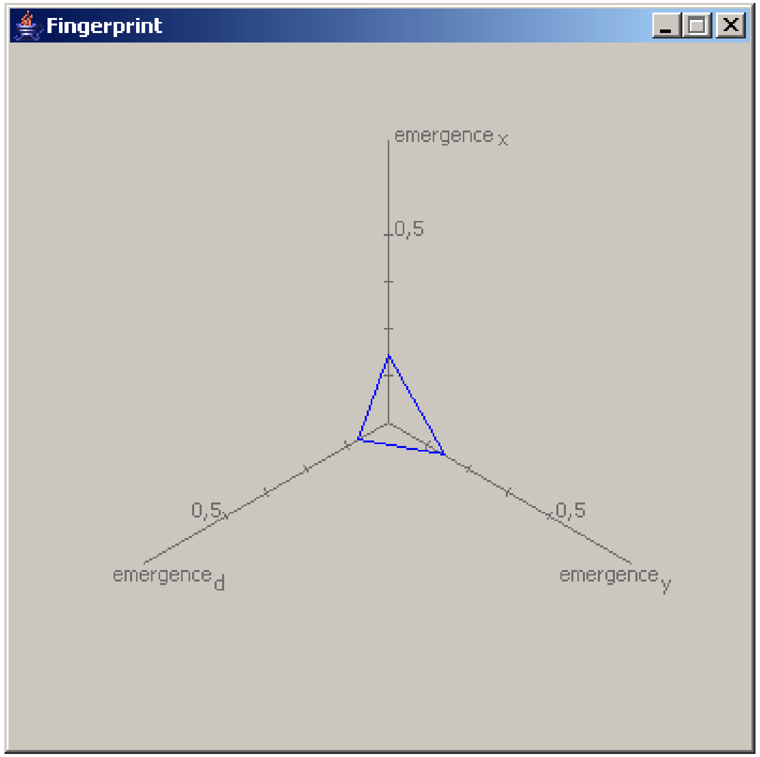
- Emergence can be calculated for a given system for different attributes.
- It can be used as an early indicator of (emergent) SO processes.
- System emergence (the total of all attribute emergence values) is not selective enough.
- More interesting: Emergence fingerprint for all relevant attributes.
- Open questions:
  - Which attributes are relevant?
  - What is positive (wanted) and negative (unwanted) emergence?
- Example: Cannibalistic behavior of chickens
  - Concentration movements toward hurt chicken: negative
  - Concentration movement toward feeding trough: neutral

- Emergence fingerprint = visualization of all (relevant) attribute emergence values of a system.
  - --> n-dimensional Kiviat graph
- Example
  - x-position, y-position, color, state, direction, intention



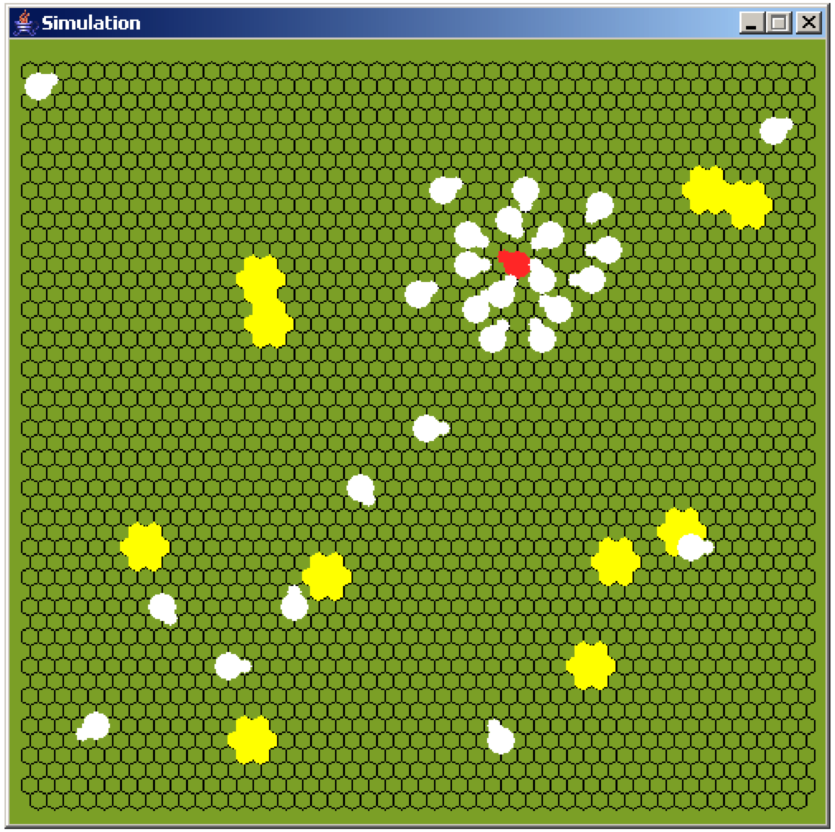
Emergence fingerprints (2)

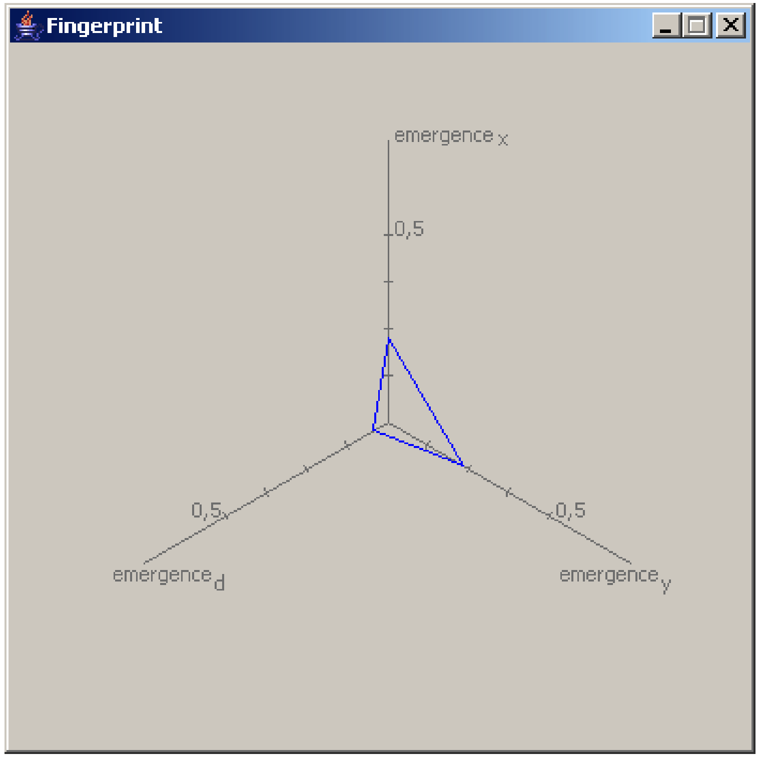




Pattern 1: $M_x = 0.181$, $M_y = 0.177$, $M_{direction} = 0.091$

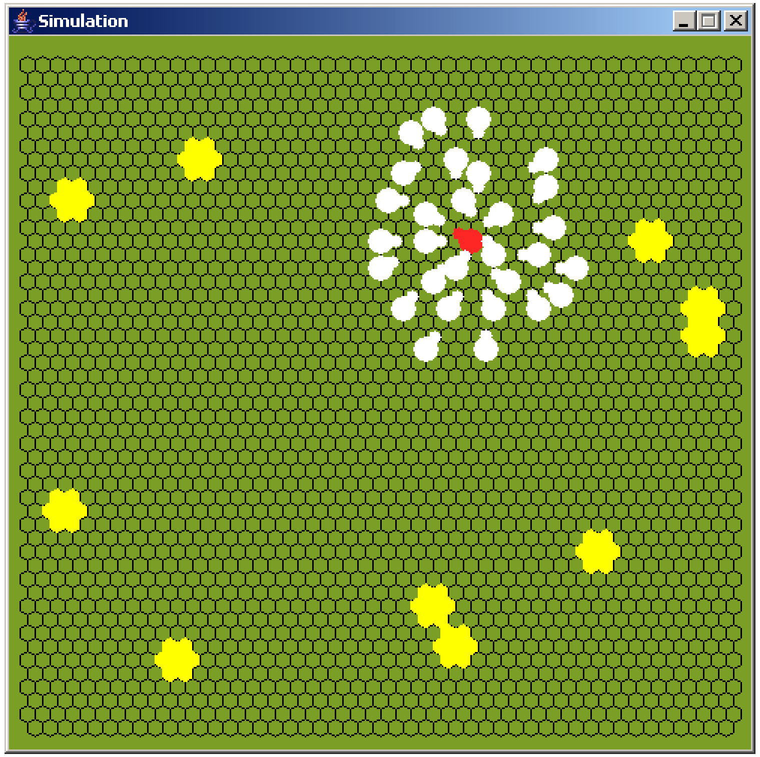
Emergence fingerprints (3)

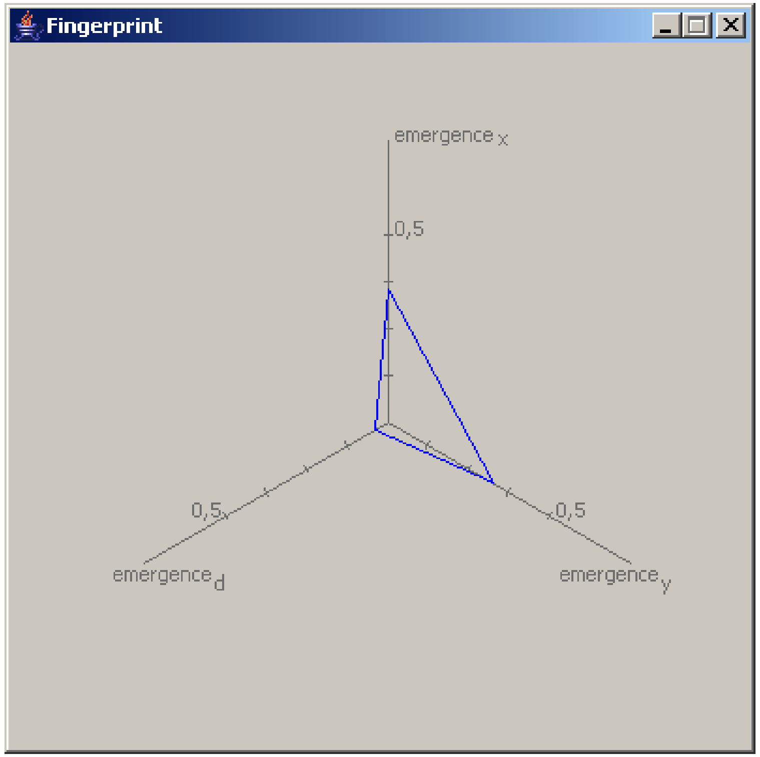




Pattern 2: $M_x = 0.226$, $M_y = 0.237$, $M_{direction} = 0.046$

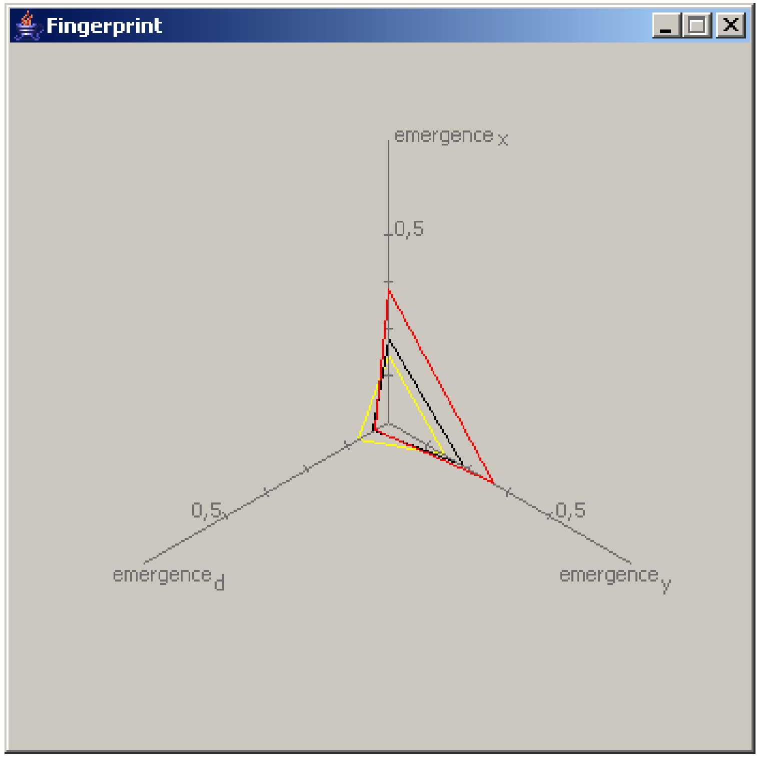
Emergence fingerprints (4)

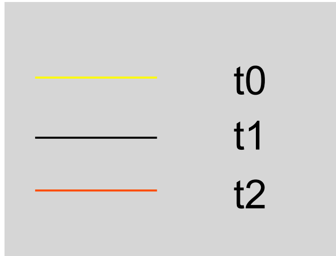




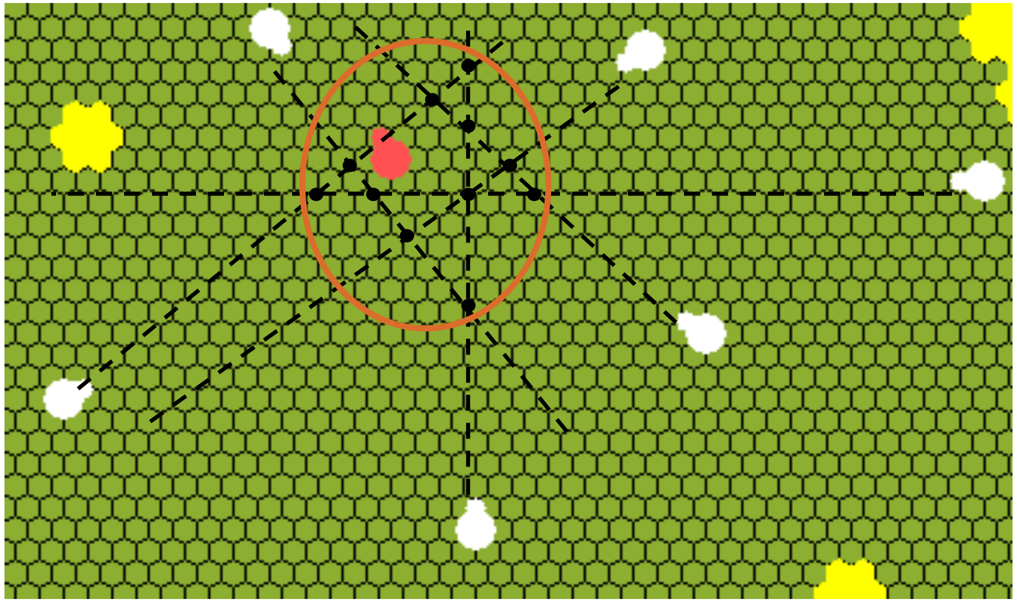
Pattern 3: $M_x = 0.359$, $M_y = 0.328$, $M_{direction} = 0.041$

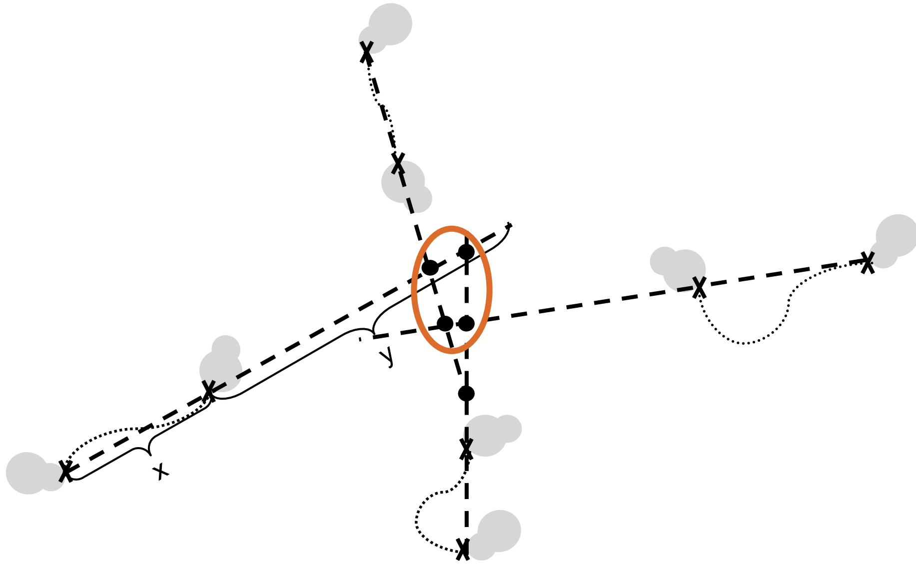
Emergence fingerprints (5)





Emergence prediction





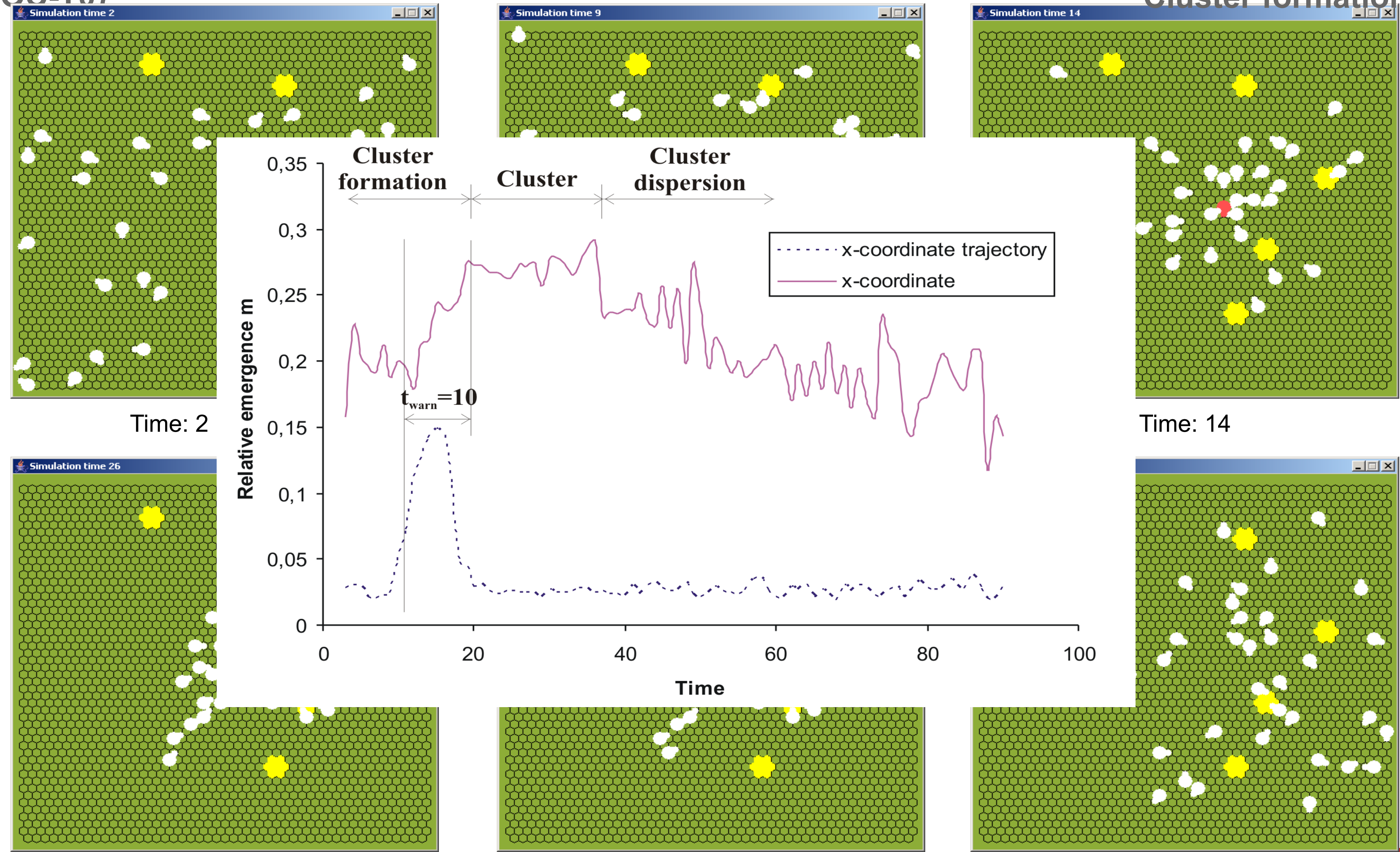
$$v = \frac{x}{\Delta t}$$

$$y = v \times \tau$$

Δt Observation period

τ Prediction period

Cluster formation



Time: 2

Time: 14

Time: 26

Time: 28

Time: 62

Summary (1)

- Verbal definitions of emergence as starting point
- Goal: Assign a high value to a system which is rated intuitively as emergent.
- Emergence is always associated with patterns (symmetry breaks) --> Structural emergence
- Patterns represent unequal distributions (“order”).
- The (non)uniformity of a distribution can be measured in terms of entropy.
- Therefore we have
  - determined entropy and
  - related it to „emergence“.

Summary (2)

- ❑ Emergence values can be practically measured.
  - Structure detectors
  - See → Observer/controller architectures
- ❑ Emergence values constitute a kind of fingerprint of characteristic system states.
- ❑ Emergence is observer dependent --> observer model, preprocessing
- ❑ Certain patterns (periodic) require preprocessing (e.g. Fourier analysis).
- ❑ Our emergence definition covers only structural emergence.
  - ❑ Is there anything else?
- ❑ This metric measures the result of a process!