Swarm Intelligence

Models and applications

Eric Bonabeau & Jean-Louis Dessalles

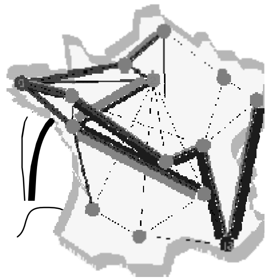
Eric Bonabeau, Marco Dorigo, Guy Théraulaz
Swarm Intelligence - Oxford University Press 1999

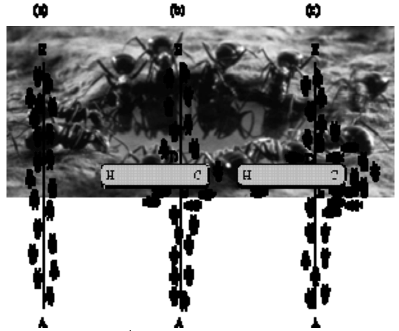
ParisTech - ENST - 2007 J-L. Dessalles

2

Swarm intelligence

- Examples
  - **Ant colony: Shortest path**





3

Biological metaphors

- Scientific objective: modelling
- Technical objective: new engineering methods

- Strong points of these metaphors:
  - ➔ *decentralization*
  - ➔ *parallelism*
  - ➔ *flexibility, adaptivity*
  - ➔ *"robustness" (failures)*
  - ➔ *auto-organization*

4

Biological metaphors

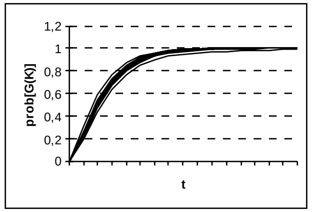
Brain	⇨	Neural networks
Evolution	⇨	Genetic Algorithms Fitness landscapes
Ants	⇨	Swarm Intelligence
Immune System	⇨	Protection of computers and networks

5

Swallows

$$\Pr[G(K+1)|G(K)] = \frac{\Pr[G(K+1) \,\&\, G(K)]}{\Pr[G(K)]}$$
$$= \frac{\Pr[G(K+1)]}{\Pr[G(K)]}$$
$$= 1 - [1 - p_1(K)]^{N-K}$$

$G(K)$ probability for a group to reach at least size K at time t



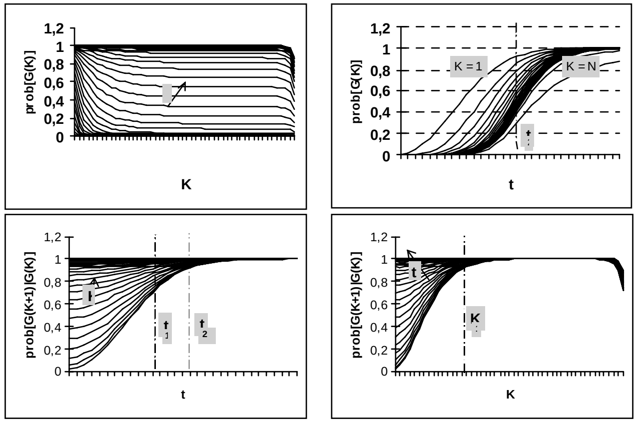
$p_1(K)$ probability for a bird to join a group of size K at time t.

$$p_1(K) = AK + Bt + C$$

6

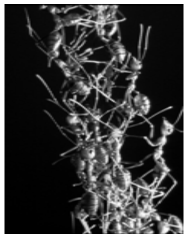
Collective Transistor

$$p_1(K) = AK^2 + Bt^2 + C$$



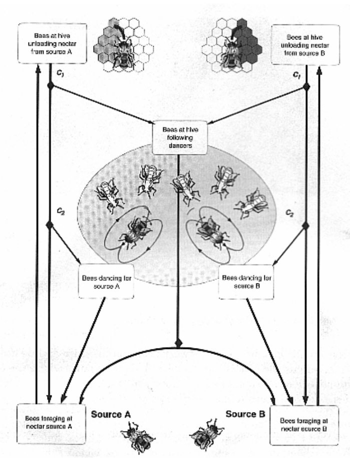
7

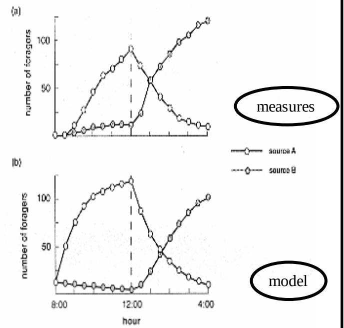
Social insects



8

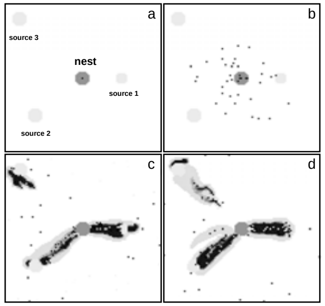
Amplification of fluctuations and "lock-in"





9

Amplification et evaporation



- The virtual colony exploits closest resources first.
- When closest food sources are exhausted, it starts to exploit farther sources.

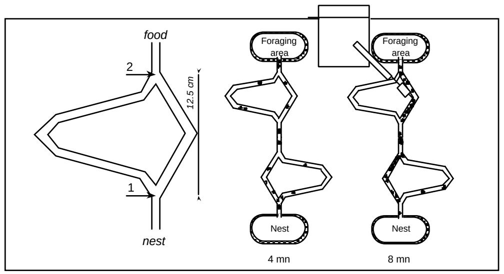
10

Four ingredients of self-organization

- Positive Feedback
- Negative Feedback
- Amplification of Fluctuations - randomness
- Reliance on multiple interactions

11

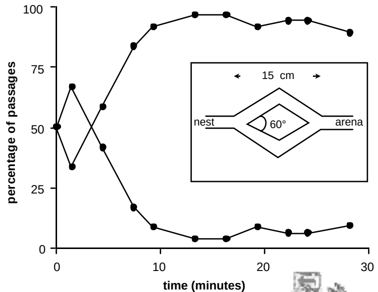
Amplification of fluctuations and optimization

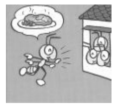


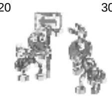
Ants collectively select the shorter path.

12

Amplification of fluctuations and "lock-in"







- The double-bridge experiment.
- On branch is almost ignored after some time.

13

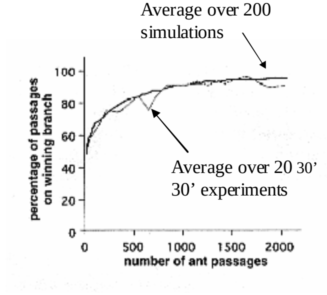
Double bridge: model

⊙ **Probability of choosing branch A**

$$P_A = \frac{(k + A_i)^n}{(k + A_i)^n + (k + B_i)^n} = 1 - P_B$$

i : number of ants crossing the bridge

A_i: number of ants having gone through branch A



p_A plot

$n \approx 2$ $k \approx 20$

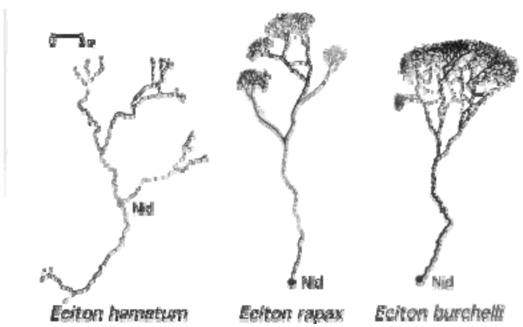
14

Properties of self-organization

- Creation of structures
  - ⊙ **Nest, foraging trails, or social organization**
- Changes resulting from the existence of multiple paths of development
  - ⊙ **Non-coordinated & coordinated phases**
- Possible coexistence of multiple stable states
  - ⊙ **Two equal food sources**

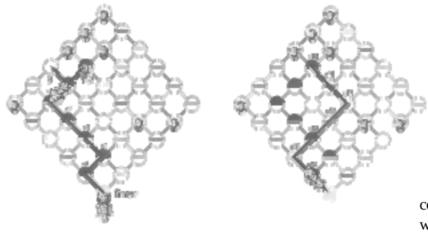
15

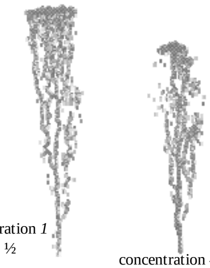
Foragement strategies



16

Foragement strategies





concentration *1* with pr. ½

concentration *400* with pr. 1/100

- When moving forward: lay down one pheromon unit (max 1000)
- When returning: 10 units (max 300)
- Evaporation: 1/30
- Probability of moving (p_m) and of choosing direction (p_l, p_r)
- Maximum 20 ants per site

$$p_m = \frac{1}{2}\left[1 + \tanh\left(\frac{\rho_l + \rho_r}{100} - 1\right)\right]$$

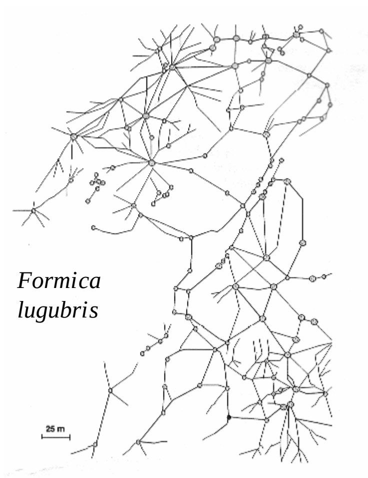
prob. of moving

$$p_l = \frac{(5 + \rho_l)^2}{(5 + \rho_l)^2 + (5 + \rho_r)^2}$$

prob. of choosing left

17

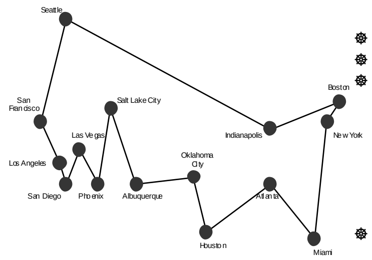
Path search in a graph



Formica lugubris

18

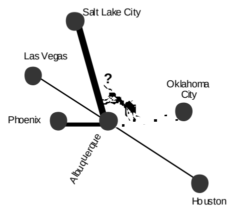
Travelling salesperson and virtual ants



- *N* cities
- distance function *d* between cities
- Find a tour, so that
  - (1) each city is visited once
  - (2) total distance is minimum
- NP-hard problem
- Classical benchmarking problem for optimization methods

19

Travelling salesperson and virtual ants



- *m* agents, each one makes a tour
- memory of visited cities
- d_{ij}= distance between city *i* and city *j*
- τ_{ij}= virtual pheromon on link (*i*,*j*)
- When in city *i*, the probability of going from city *i* to city *j* is proportional to $(\tau_{ij})^{\alpha} (d_{ij})^{-\beta}$
- At the end of a tour of length *L*, each agent reinforces the links it went through with a quantity proportional to *1/L*
- Virtual pheromon evaporates : $\tau \rightarrow (1-\rho)\ \tau$

20

Ant system

- Not only do the ants find a very good solution to the problem, they also maintain a pool of alternate solutions.
- In case a city or a link is added or disappears, the system can quickly reorganize itself and find a good solution to the new situation.

$$p_{ij}^k(t) = \frac{[\tau_{ij}(t)]^{\alpha} \cdot [\eta_{ij}(t)]^{\beta}}{\sum_{l \in J_i^k} [\tau_{il}(t)]^{\alpha} \cdot [\eta_{il}(t)]^{\beta}} \qquad \eta_{ij} = 1/d_{ij}$$

$$\tau_{ij}(t) \leftarrow (1-\rho) \cdot \tau_{ij}(t) + \Delta\tau_{ij}(t) \qquad \Delta\tau_{ij}^k(t) = \begin{cases} Q/L^k(t) & \text{si } (i,j) \in T^k(t) \\ 0 & \text{si } (i,j) \notin T^k(t) \end{cases}$$

Other applications

The same method may be applied to any allocation problem

- Traveling salesman problem
- Quadratic assignment problem
- Job-shop scheduling
- Graph coloring
- Vehicle scheduling

AS-TSP : Traveling salesman problem

	Best tour	Average	Std. Dev.
Simulated Annealing	422	459.8	25.1
Tabu search	420	420.6	1.5
AS-TSP	420	420.4	1.3

QAP: quadratic assignment problem

- Allocate n activities to n locations. $\pi(i)$: activity assigned to i.
- Find a permutation that minimizes a cost function by taking into account the flow of exchanges beetween activities

$$\pi_{opt} = \arg \min_{\pi \in \Pi(n)} C(\pi) \qquad C(\pi) = \sum_{i,j=1}^{n} d_{ij} f_{\pi(i)\pi(j)}$$

	Nugent (7)	Nugent (12)	Nugent (15)	Nugent (20)	Nugent (30)	Elshafei (19)	Krarup (30)
SA	148	578	1150	2570	6128	17937024	89800
TS	148	578	1150	2570	6124	17212548	90090
GA	148	588	1160	2688	6784	17640584	108830
ES	148	598	1168	2654	6308	19600212	97880
SC	148	578	1150	2570	6154	17212548	88900
AS-QAP	**148**	**578**	**1150**	**2598**	**6232**	**18122850**	**92490**
AS-LS	148	578	1150	2570	6146	17212548	89300
AS-SA	148	578	1150	2570	6128	17212548	88900

QAP: quadratic assignment problem

Potential Vectors

$$d_i = \sum_{j=1}^{n} d_{ij} \qquad f_h = \sum_{k=1}^{n} f_{hk} \qquad E = \overline{d} \cdot \overline{f}^T$$

- An initial solution is constructed using the minimax rule: The reminding location with lowest potential receives the reminding activity with highest potential.
- The ant algorithm is applied: it goes through locations with increasing potential, with:

$$\eta_{ij} = d_i \cdot f_j$$

$$\Delta\tau_{ij}^k = Q / C^k(t) \text{ if ant } k \text{ chose allocation } (i, j)$$

25

Dynamics

☸ Many problems are by nature dynamic. Their formulation varies as time goes, either because the system's internal characteristics change, or because external conditions change.

☸ Variation time scale is essential. It is sometimes impossible to apply an exhaustive method. Optimization must be dynamic.

☸ Variations may be so rapid that optimization becomes less important than fulfulling the task.

26

Robustness and flexibility

☸ *Robustness* : A system is robust if it keeps functioning efficiently even if some of its constituent parts fail.

☸ *Flexibility* : A système is said to be flexible if it can efficiently function when external conditions change.

27

Robustness and flexibility

☸ *Robustness* : For example, an assembly line is robust if production continues when a machine fails. Robustness degree : How many machines may break down without (too) affecting production ?

☸ *Flexibility* : an assembly line is flexible if it can react to changing demands. Degree of flexibility : What is the reaction time, and what amount of fluctuation can it tolerate?

28

Optimization with artificial ants

Why does it work at all?

☸ Fundamental principle: reinforcement of partial solutions and global dissipation. This principle presuppose that the problem be structured (ex : ants perform well on structured instances of QAP).

☸ Other important principle: keep a distributed trace of past exploration. Optimization efficiency and reaction to changing conditions are improved, because of the distributed memory of alternate solutions.

29

Similar approaches

- Neural networks
- Population-based incremental learning PBIL (Baluja & Caruana 1995)
- Bit-based simulated crossover (Syswerda 1993)
- Mutual Information Maximization for Input Clustering MIMIC (De Bonet et al. 1997)
- Bayesian Networks

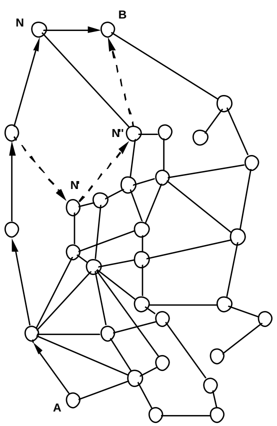
30

Routing in telephone networks

- Routing : Device that processes the next direction of a message at a node of the network
- Messages should reach their destination
- Time needed to go from the source to the destination must be kept minimal
- Characteristics of the trafic change constantly: routing must adapt

31

Wy routing ?

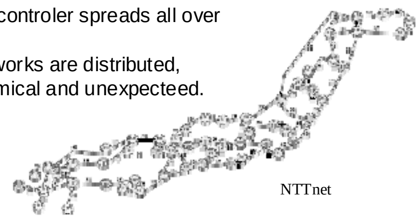


If node A sends a message to node B, the message has to go through a set of intermediate nodes because A and B are not directly connected. One possible shortest path for the message is the one indicated by thick lines and arrows, which takes the message from A to B in 5 steps. If, however, node N breaks down or is highly congested, the message needs to be rerouted dynamically toward a slightly longer route that goes through nodes N' and N''. Although it now takes 6 hops for the message to be transmitted from A to B, the actual transmission time will be reduced and the message will be less likely to be lost.

32

Routing

- Switching nodes hold routing tables that direct messages to other nodes depending on their final destination.
- Routing tables are regularly updated by a centralized mechanism:
  - ➔ Requires centralization and increases traffic
  - ➔ Maladpated to large networks
  - ➔ Failure at the central controler spreads all over the network
  - ➔ Communications networks are distributed, spatially extented, dynamical and unexpecteed.

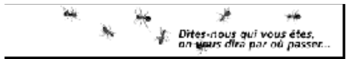


NTTnet

33

Ants in the network!

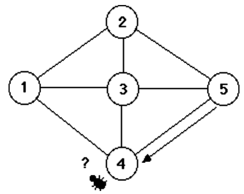
- Ant agents are launched in the network.
- An agent updates routing tables *by considering its source as a destination*.
  - "If you are going to my source, go first to the node I am coming from (if I am 'young' enough)"
  - Or "Don't go there (if I am old)".
- It influence diminishes with "age".
- Agents are made artificially older at overload nodes.



34

Ants in the network!

Example of network and of routing table.



Destination nodes

Neighbor nodes	1	2	3	5
1	0.8	0.3	0.1	0.1
3	0.1	0.4	0.8	0.1
5	0.1	0.3	0.1	0.8

Messages, contrary to ants, travel in the network deterministically, always following highest probability.

Probability of directions for an ant going to node 2.

Probabilities updated by an ant coming from node 5.

demo

35

Ants in the network!

Schoonderwoerd et al. (1996)

$r^i_{n,d}(t)$ Probability, at node *i*, when heading to node *d*, of choosing *n* as next node.

$$r^i_{i-1,s}(t+1) = \frac{r^i_{i-1,s}(t) + \delta r}{1+\delta r}$$

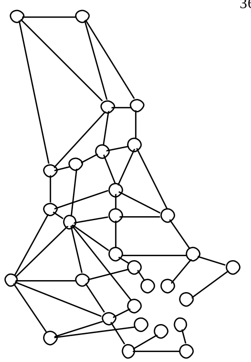
$$r^i_{n,s}(t+1) = \frac{r^i_{n,s}(t)}{1+\delta r}, \quad n \neq i-1$$

$$\delta r = \frac{a}{T} + b$$ T: ant's age

$$D = c \cdot e^{-d \cdot S}$$ D: delay;
S: remaining capacity of the node

36

Model network



Model network used in simulations: BT interconnection network.

(Little game: where do you think London is?)

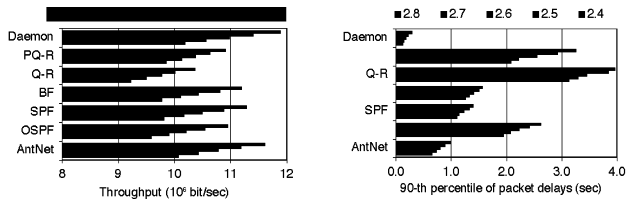
Performance of ABC (ant-based control):

	Average call failures	Std. Dev.
Shortest path	12.57	2.16
Mobile agents	9.19	0.78
Improved mobile agents	4.22	0.77
ABC without noise	1.79	0.54
ABC with noise	1.99	0.54

37

Results (with AntNet)

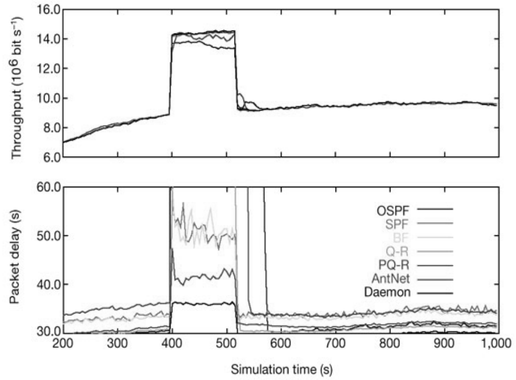
Poisson traffic on NSFNET, various interarrival times.



OSPF: Open Shortest Path First (current Internet routing algorithm),
SPF: Shortest Path First, BF: Bellmann-Ford, [P]QR: [Predictive] Q-Routing

38

Results (with AntNet)

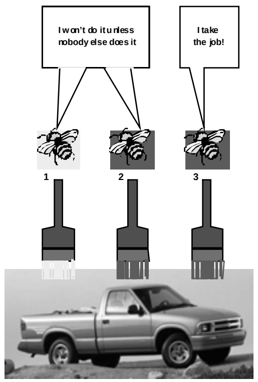


Hot spot superim-posed to Poisson traffic.

Moving average over 10s.

39

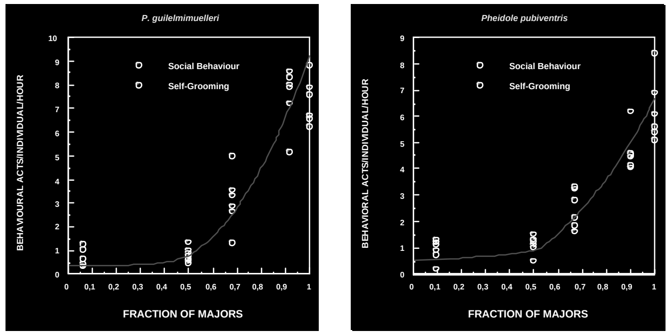
From division of labor to scheduling



- Scheduling technique inspired by task allocation in a honeybee colony: individual bees are specialized in certain tasks, which depend on their age, but they can perform other tasks if needed. For example, a nurse bee can become a forager bee if there is not enough food coming into the hive.
- Our assumption is that a bee performs the tasks for which it is specialized unless it perceives that other tasks badly need to be performed.
- To allocate trucks coming out of an assembly line to paint booths in a truck factory, each paint booth is considered an artificial bee specialized in one color. But if needed, the paint booth can change its color (though it's costly).
- The system minimizes paint changes and can cope with glitches.

40

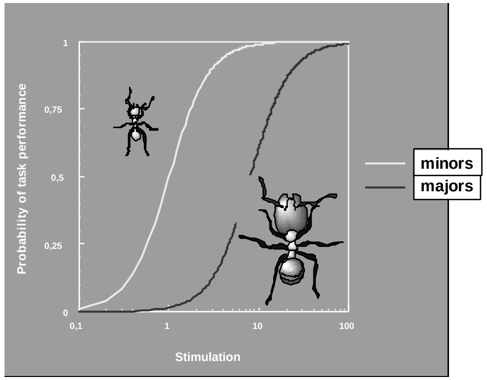
Colony-level flexibility in two ant species (*Pheidole*)



Augmentation de la proportion de *majors* par retrait expérimental des *minors*

41

Response curves: differential thresholds



42

Model

Probability of performing the task for a stimulus *s*:

$$T_\theta(s) = \frac{s^n}{s^n + \theta^n}$$

$$T_\theta(s) = 1 - e^{-s/\theta}$$

ρ= prob. of encountering an item

$$P(N) = 1 - (1-\rho)^N = 1 - e^{N \ln(1-\rho)}$$

n = 2

$\theta_{\text{(task 1, major)}} = 8$

$\theta_{\text{(task 1, minor)}} = 1$

$P_{\text{(active->inactive)}} = 0.2$ (per time step)

$\text{stimulus}_{(t+1)} = \text{stimulus}_{(t)} + (1 - (3 \frac{N_{\text{(active)}}}{N_{\text{(population)}}}))$

plot *seuils*

43

Threshold reinforcement

Fixed thresholds cannot account for genesis of specialization in non-polymorphic species.

Although tasks are eventually completed when the system is perturbed, there may be an irreversible degradation of the system's performance: stimulus intensity remains high.

⇨ Threshold reinforcement: the more an agent performs a task, the lower its response threshold. New specialists can be generated in response to perturbations.

44

Threshold reinforcement: application

Mail retrieval in a city:

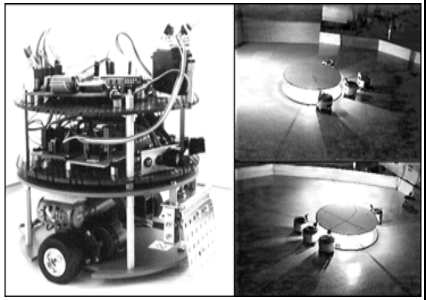
- N agents
- City divided into zones
- Each agent has response thresholds for all zones
- Agent responds to demand from a zone when stimulus exceeds threshold
- Current working zone's threshold is reinforced, as well as neighboring zones' thresholds. All other thresholds decay

⇨ Specialization and robustness

45

Cooperative transport

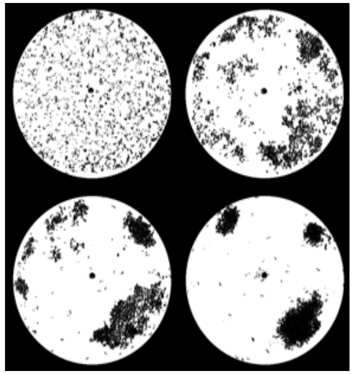
- Observed in several ant species: a single ant cannot retrieve a large prey, nestmates are recruited to help. Then, during an initial period of up to several minutes, the ants change position and alignment around the prey without making progress, until eventually the prey can be moved toward the nest.
- Ron Kube and Hong Zhang have reproduced this emergent coordination with a swarm of very simple robots. Videotaped experiments worth viewing at http://www.cs.ualberta.ca/~kube/.
- Not the most efficient way of pushing a box. But, because of the simplicity of the robots, promising in the perspective of miniaturization and low-cost robotics.



Box pushing

46

Cemetery formation in *Messor sancta*



- Workers form piles of their dead nestmates' corpses –literally cemeteries– to clean up their nests.
- If corpses are randomly distributed in space at the beginning of the experiment, the workers form clusters within a few hours (figure shows the initial state with 1500 corpses, 2 hours, 6 hours, and 26 hours after the beginning of the experiment).
- Small clusters of items grow by attracting workers to deposit more items.
- Brood sorting follows same type of logic: an ant picks up and drops an item according to the number of similar surrounding items.

47

Clustering in ants

- An isolated item is more likely to be picked up by an unladen agent:

$$P_p=[k_1/(k_1+f)]^2$$

where f=density of items in neighborhood

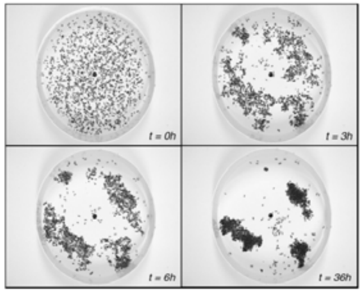
- A laden agent is more likely to drop an item next to other items:

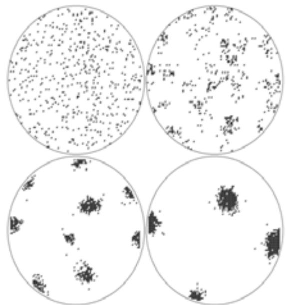
$$P_d=[f/(k_2+f)]^2$$

plot *probabilities*

48

Cemetery formation in *Messor sancta*





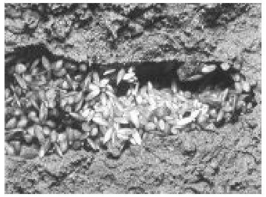
49

From clustering to sorting

- The same principle can be applied to sort items of several types (i=1,...,n).
- f is replaced by f_i, the fraction of type i items in the agent's neighborhood:

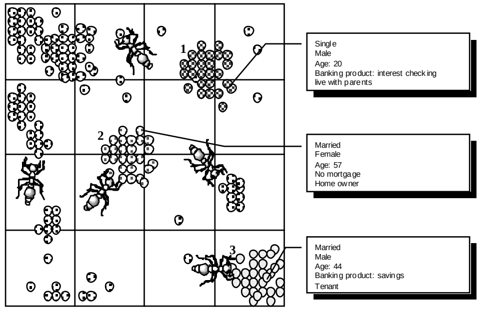
$$P_p(i)=[k_1/(k_1+f_i)]^2$$

$$P_d(i)=[f/(k_2+f_i)]^2$$



50

From brood sorting to data analysis



- Artificial ants move around and pick up and drop "clients" according to how many similar clients there are in the neighborhood.
- The measure of how similar two clients are is based on a natural distance for each of the attributes. For example, for attributes such as marital status or gender, a similarity value of 1 is assigned to pairs having the same value of the attribute, and a value of 0 to pairs with different values. For age, the smaller the age difference the higher the similarity.
- Emergent clusters obtained and visualized.

51

From sorting to data analysis

- If items are described by real-valued attributes (points in R^n), the same principle can still be applied: f is now replaced by a normalized distance between the item carried by the agent and items in the agent's neighborhood.

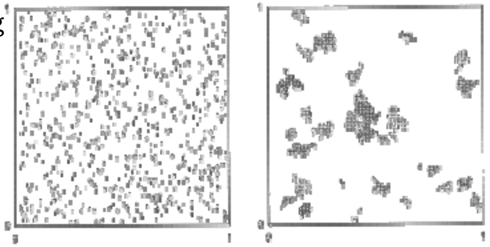
$$f(o_i) = \begin{cases} \frac{1}{s^2} \sum_{o_j \in Neigh_{(s \times s)}(r)} \left[1 - \frac{d(o_i, o_j)}{\alpha}\right] & \text{if } f > 0 \\ 0 & \text{otherwise} \end{cases}$$

⇨ Items will end up being next to items with close attributes.

α contrôle la discrimination entre objets

52

Ant clustering



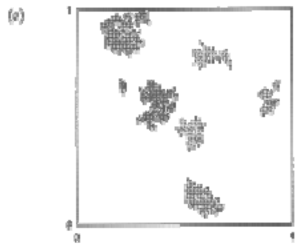
$k_1 = 0.1$

$k_2 = 0.15$

$\alpha = 0.5$

$s^2 = 9$

$t = 500.000 - -1.000.000$



53

Graph representation

- Same method can also be applied to graph drawing. Complex networks of relationships arise in many contexts and can often be represented as graphs. Drawing a graph in the plane facilitates interpretation by observer.
- Vertices in a graph have attributes: the vertices they are connected to. A good distance between vertices is the number of adjacent vertices they have in common.
- Example: random graphs with clusters.

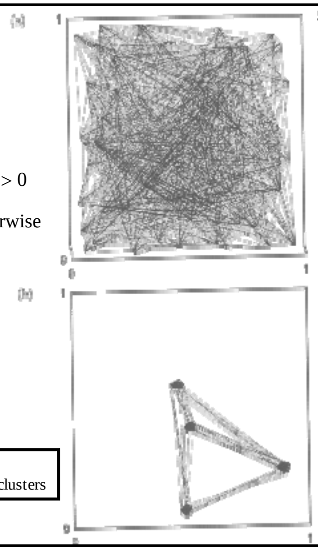
déplacement

54

Graph representation

$$f(v_i) = \begin{cases} \frac{1}{s^2} \sum_{v_j \in Neigh_{(s \times s)}(r)} \left[1 - \frac{d(v_i, v_j)}{\alpha}\right] & \text{if } f > 0 \\ 0 & \text{otherwise} \end{cases}$$

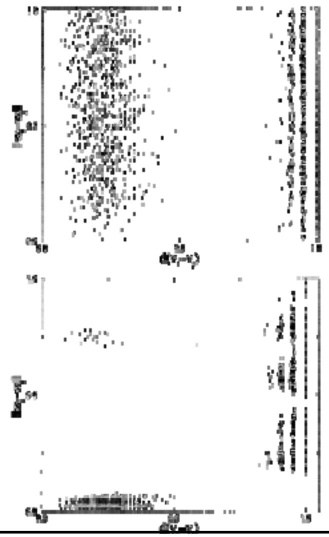
$$d(v_i, v_j) = \frac{\left|\Delta\left(\rho(v_i), \rho(v_j)\right)\right|}{\left|\rho(v_i)\right| + \left|\rho(v_j)\right|}$$



partitioning of a random graph
$\Gamma(25,4,0.8,0.01)$ with 25 vertices and 4 clusters

55

w_i = position of v_i on the plane



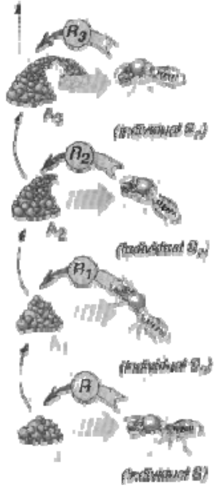
56

Morphogenesis

Turing Démo d'Arcy

57

Stigmergy



58

Reaction-diffusion model of the royal chamber construction

$H(r,t)$ Pheromon concentration in *r* at time *t*

P quantity of active material

C density of laden termites

Φ laden termite entering flow

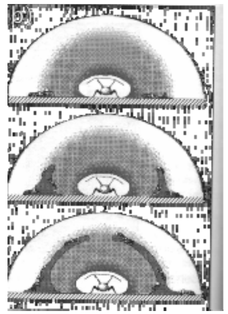
$T(x,y)=e^{-\left[((x-x_0)/\lambda_x)^2+((y-y_0)/\lambda_y)^2\right]}$ template

$$\partial_t H = k_2 P - k_4 H + D_H \nabla^2 H$$

$$\partial_t C = \Phi - k_1 C + D_C \nabla^2 C - \gamma\nabla(C\nabla H) - \nu\nabla(C\nabla T)$$

$$\partial_t P = k_1 C - k_2 P$$

Macrotermes subhyalinus



59

Self-organization in the presence of templates

U : density of unladen ants

L : density of laden ants

S : grain density

$P(r)$: influence of the template to pick grain

$P(r)F(S)US$: transition rate $U \rightarrow L$

$F(S)$: $(g_1+g_2S)^{-1}$ perception of grain density by ants

$D(r)G(S)L(1-S/K)$: transition rate $L \rightarrow U$

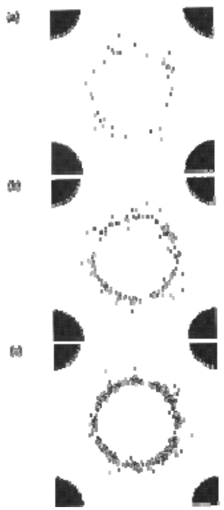
$D(r)$: influence of the template to drop grain

K : max. density

$G(S)$: (g_1+g_2S) ant's perception

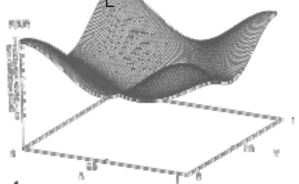
$$\partial_t S = D(r)G(S)L\left(1-\frac{S}{K}\right) - P(r)F(S)SU$$

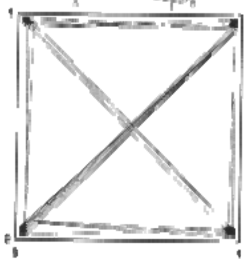
Leptothorax albipennis



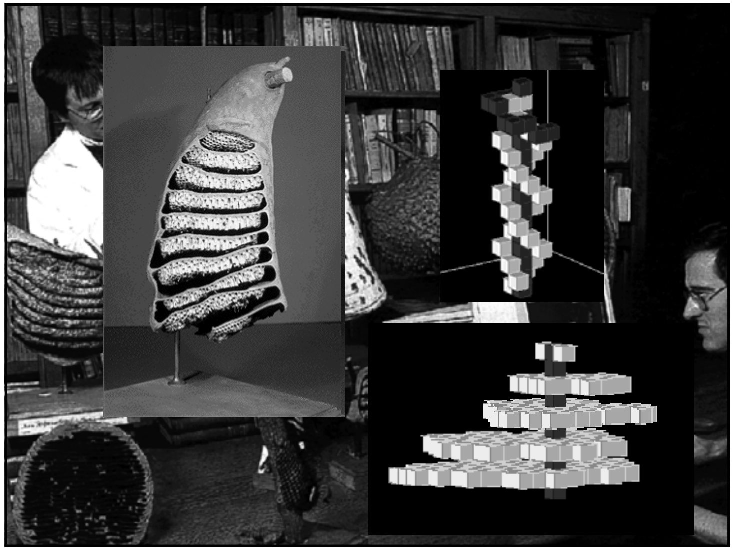
60

$$P_t(x,y)=a\left[e^{-\frac{x^2+y^2}{\sigma^2}}+e^{-\frac{(x-1)^2+y^2}{\sigma^2}}+e^{-\frac{(x-1)^2+(y-1)^2}{\sigma^2}}+e^{-\frac{x^2+(y-1)^2}{\sigma^2}}-b\right]$$



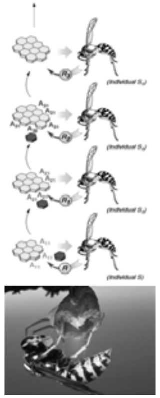


$p_p(o_i)=[k_1/(k_1+f(o_i))]^2+\alpha(1-P_t)$

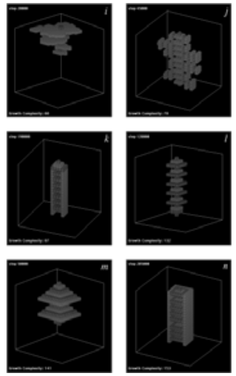


62

Wasp nest building and self-assembly



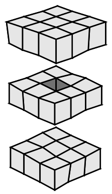
From a model of wasp nest building to self-assembly



63

Model of Building in Social Wasps

- Agents move randomly on a 3D grid of sites.
- An agent deposits a brick every time it finds a stimulating configuration.
- Rule table contains all such configurations. A rule table defines an algorithm.
- Rule space is very large.



64

Simulation model of wasp building

- Most algorithms generate structureless shapes.
- But some produce "structured" architectures.
- Structured architectures:
  - Usually modular
  - Most complex patterns have large modules
  - Produced by specific algorithms
  - Convergence to similar shape in all runs
  - Compact
  - Take time to generate

Stimulating configurations corresponding to different building stages must not overlap

65

Genetic algorithm to explore rule space

Some of the characteristics of "structured" architectures can be formalized (graph associated with the building process) and quantified.

Quantification is useful to define a fitness function. Heuristic fitness correlates well with observers' notion of structure. A GA has been run with this fitness.