A comparison of stochastic and analytical models for cell migration during early embryo somitogenesis

Christian A. Yates

Centre for Mathematical Biology
University of Oxford

UK Systems Biology inter-DTC conference,
University of Manchester Systems Biology DTC

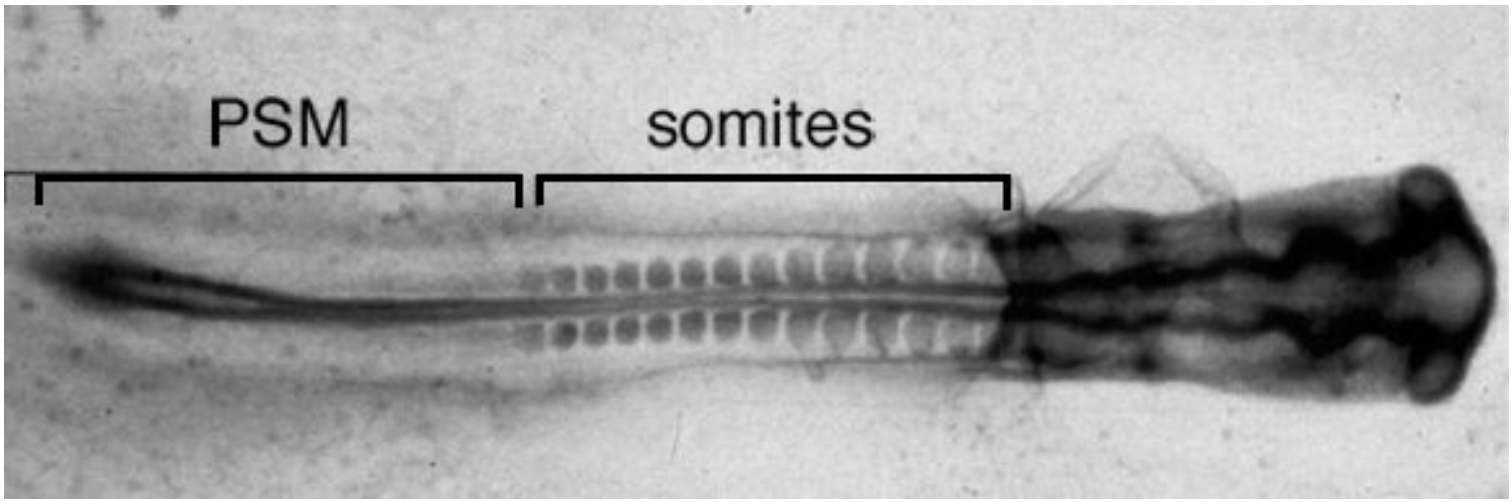
June 18th, 2009

Outline

- What is Somitogenesis?
- Why do we want to model it?
- Other application areas
- Modelling approaches for cell migration: stochastic vs continuum.
- Equivalence between the two approaches on the stationary domain: theory and simulation.
- Equivalence on the growing domain: simulation and theory.
- Where next?

What is Somitogenesis?

- Sequential formation of somites!
- Segmented mesodermal tissue either side of the neural tube along the A-P (head-tail) axis.
- Precursors to dermis, skeletal muscles and axial skeleton.
- Form at anterior (head) end of PSM and proceed posteriorly (towards the tail).
- Common to all vertebrates

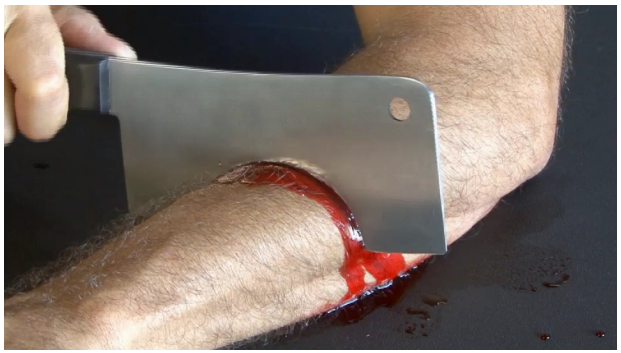


Chick embryo displaying somites after 40-45 hours of incubation.

Why do we want to model cell migration?

Cell migration is involved in many biological processes

- Diseases caused by erroneous cell migration in somitogenesis: Jarcho-Levin syndrome and congenital scoliosis.
- Cell migration also contributes to tumour invasion and metastasis.
- In later life cell migration is responsible, in part, for:
  - Wound healing
  - Tissue repair
  - Some immune responses



Two Contrasting Modelling Approaches

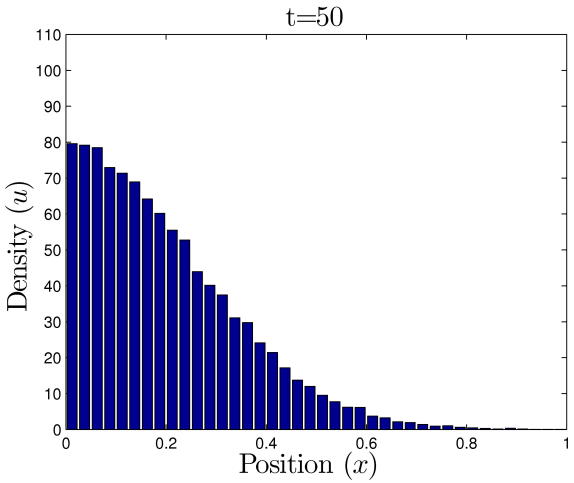
Given initial cell density across the domain and boundary conditions how do we find cell density at later times?

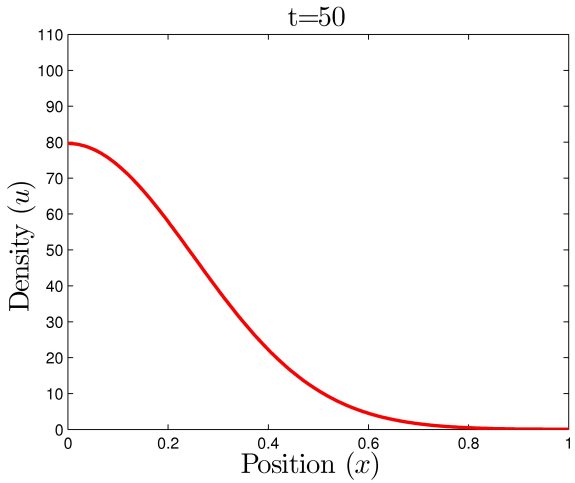
Stochastic/Discrete Approach

- Gillespie Algorithm
- $\alpha_0 = \sum_{i=1}^{k} \alpha_i^+ + \sum_{i=1}^{k} \alpha_i^-$
- $\tau = \frac{1}{\alpha_0} \ln\left(\frac{1}{rand_1}\right)$

Deterministic/Continuum Approach

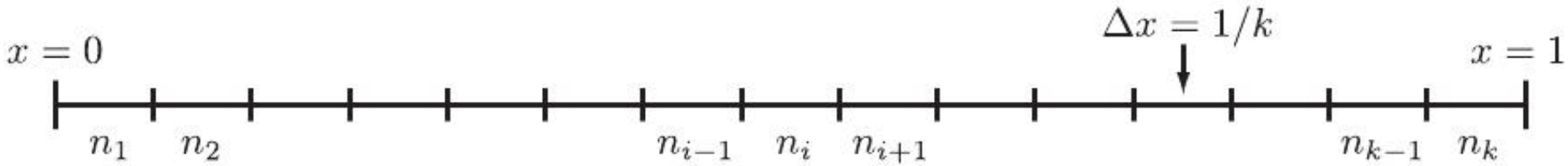
- Partial Differential Equations
- $\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2},$
- $(x, t) \in [0, 1] \times [0, \infty)$





Aside - What is a Gillespie Algorithm? - Background

- Assume $k \geq 1$ (chemical) species $\{N_1 \dots N_k\}$.
- Classify a set of $M \geq 1$ reactions (chemical or otherwise) $\{R_1 \dots R_M\}$.
- Associate with each reaction, R_j a 'propensity function' $a_j(\mathbf{N})$, which describes the probability one R_j reaction will happen in the next time interval $[t + dt]$.



In our model:

- k is the number of boxes we split the domain into.
- $\{N_1 \dots N_k\}$ represents the number of cells in each box.
- $\{R_1 \dots R_M\}$ are the reactions of a cell moving left or right out of the box.
- Hence $M = 2k$.
- The propensity of each reaction is proportional to the number of cells in the corresponding box.

Aside - What is a Gillespie Algorithm? - The Algorithm

The Algorithm

1. Initialize (i.e., set initial numbers of species, set t=0).
2. Calculate the propensity functions, a_j, for $j = 1 \dots M$.
3. Generate a uniform random number, r_1, and choose a time τ for the next reaction (using formula $\tau = \frac{1}{\sum_{j=1}^{N} a_j} \log(\frac{1}{r_1})$).
4. Generate a second uniform random number (r_2) and choose which of the reactions will happen (with probability proportional to their propensity functions). i.e. find i s.t. $\sum_{j=1}^{i-1} a_j \leq r_2 \leq \sum_{j=1}^{i} a_j$.
5. Change the number of species to reflect execution of the chosen reaction.
6. update $t \leftarrow t + \tau$.
7. Go to step 2.

Advantages and Disadvantages

Stochastic/Discrete Approach

- Statistically accurate with low numbers of cells.
- Allows for the incorporation of noise.

BUT

- Slow.
- Different results each time.

Deterministic/Continuum Approach

- Fast.
- Consistent results.
- Good results with large cell numbers.

BUT

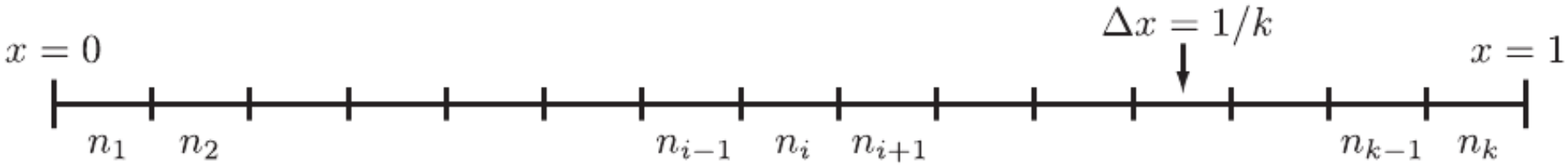
- Statistically innaccurate for low cell numbers.
- Unrealistic as no noise effects.




Stationary domain - Stochastic Individual Model... (1)

$\mathbf{N}(t) = [N_1(t), N_2(t) \dots N_i(t) \dots N_k(t)]$.

$P(\mathbf{n}, t) = P(\mathbf{N} = \mathbf{n}, t)$.



Define creation and annihilation operators A_i^c and A_i^a: $\mathbb{R}^k \to \mathbb{R}^k$ and projection operators $B_i : \mathbb{R}^k \to \mathbb{R}$, $i = 1, 2, \dots, k$

$A_i^c : [n_1, \dots, n_i, \dots, n_k] \to [n_1, \dots, n_i + 1, \dots, n_k]$,

$A_i^a : [n_1, \dots, n_i, \dots, n_k] \to [n_1, \dots, n_i - 1, \dots, n_k]$,

$B_i : [n_1, \dots, n_i, \dots, n_k] \to n_i$.

Stationary domain - Stochastic Individual Model... (2)

Master Equation

$$
\begin{aligned}
P(\mathbf{n},\mathbf{s},t+\delta t) &= \sum_{i=1}^{k-1} d\left\{[B_i A_{i+1}^a A_i^c \mathbf{n}] P(A_{i+1}^a A_i^c \mathbf{n},\mathbf{s},t)\right\}\delta t \\
&+ \sum_{i=2}^{k} d\left\{[B_i A_{i-1}^a A_i^c \mathbf{n}] P(A_{i-1}^a A_i^c \mathbf{n},\mathbf{s},t)\right\}\delta t \\
&+ P(\mathbf{n},\mathbf{s},t) - \sum_{i=1}^{k-1}\left\{[B_i \mathbf{n}] P(\mathbf{n},\mathbf{s},t)\right\}\delta t \\
&- \sum_{i=2}^{k}\left\{[B_i \mathbf{n}] P(\mathbf{n},\mathbf{s},t)\right\}\delta t
\end{aligned}
$$

... To continuous population model

Define $\langle \mathbf{n} \rangle = [\langle n_1 \rangle, \ldots, \langle n_k \rangle] = \sum_{\mathbf{n}} \mathbf{n} P(\mathbf{n}, \mathbf{s}, t)$ to be the vector of stochastic means. Then we can show

$$\begin{aligned}
\frac{\partial \langle n_1 \rangle}{\partial t} &= d\langle n_2 \rangle - d\langle n_1 \rangle, \\
\frac{\partial \langle n_i \rangle}{\partial t} &= d\langle n_{i-1} \rangle - 2d\langle n_i \rangle + d\langle n_{i+1} \rangle, \quad i = 2 \ldots k-1 \\
\frac{\partial \langle n_k \rangle}{\partial t} &= d\langle n_{k-1} \rangle - d\langle n_k \rangle.
\end{aligned}$$

This is clearly a discretisation of $\frac{\partial \langle \mathbf{n} \rangle}{\partial t} = D \frac{\partial^2 \langle \mathbf{n} \rangle}{\partial x^2}, \quad D = d(\Delta x)^2$

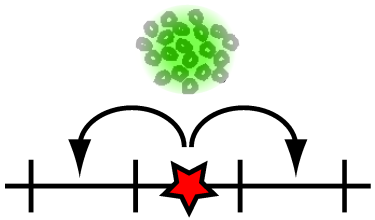
A numerical comparison of the two models of basic diffusion

(Loading movie...)

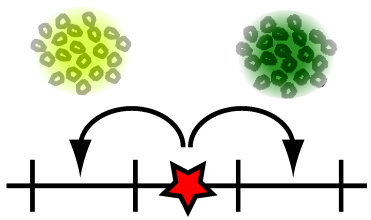
Histogram represents the average over 20 stochastic simulations and the red curve represnts the solution of the corresponding PDE.

Different sorts of signal sensing

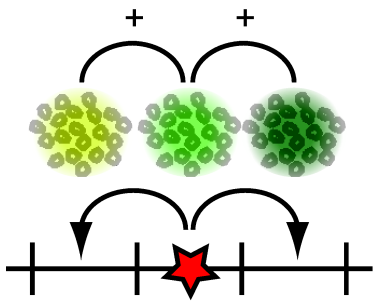
- Introduce a signal molecule concentration across the domain (i.e. s(x)=exp(-x)).
- Local - constrain cells to sense the strength of the signalling chemical only at their current site,
  $T^- = T^+ = f(s_i)$.
- Non-local - cells sense the strength of the signalling profile over a wider region, but must disregard information at the current site, i,
  $T^\pm = f(s_{i\pm1})$.
- Average - cells use information from the current position and adjacent positions,
  $T^\pm = f^l(s_i) + f^n(s_{i\pm1})$.



Local sensing



Non-local sensing



Average sensing

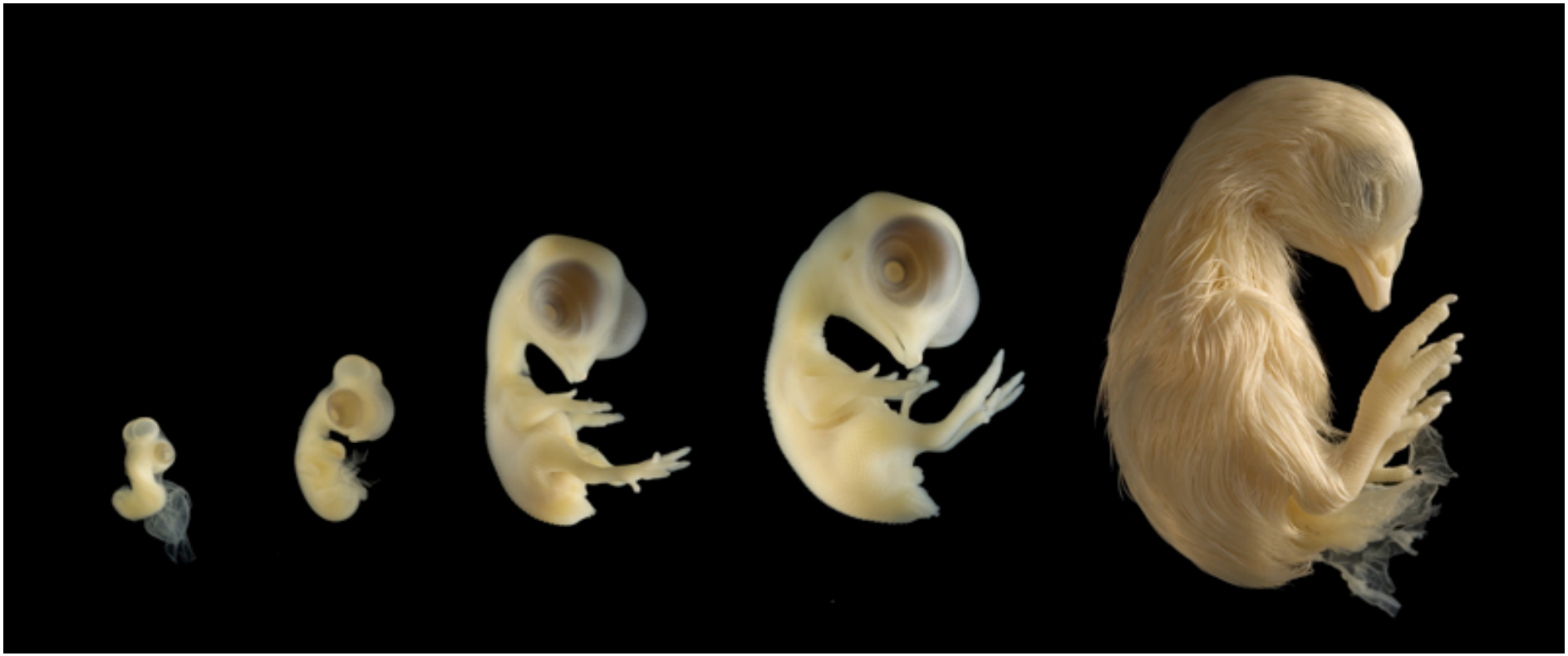
A numerical comparison of the two models of local signal sensing

(Loading movie...)

Histogram represents the average over 20 stochastic simulations and the red curve represnts the solution of the corresponding PDE.

Incorporating Domain Growth

In reality, during embryogenesis and somitogeneis cell migration occurs on a growing domain.



We need to incorporate domain growth into out model of cell migration in both the analytical and stochastic models.

Analytical derivation of Domain growth PDE

How to solve a PDE on a growing domain

- Apply Reynold's transport theorem:
  $\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{a}u) = D\nabla^2 u + R(u)$, $\mathbf{x} \in \Omega(t)$, $t \in [0, \infty)$, where $\Omega(t)$ is the now time-dependent growing domain.

- Determine the flow, a:
  $a = \frac{dx}{dt} = s(u)x$,
  where s is the (possibly density dependent) strain.

- Convert to Lagrangian (normalised) coordinates to solve the PDE.

- Transform the solution back to the growing domain.

A numerical comparison of the two models of non-local sensing on an exponentially growing domain

(Loading movie...)

Histogram represents the average over 40 stochastic simulations and the red curve represnts the solution of the corresponding PDE.

Other types of domain growth

Other more complicated types of growth

- Density-dependent domain growth:
  - Linear dependence
    $s(u) = r \times u$.
  - Quadratic dependence
    $s(u) = r \times u^2$
  - Inverse dependence
    $s(u) = \frac{r}{r+u}$
  - Many more to explore
    $s(u) =$???
- Linear domain growth
- Logistic domain growth

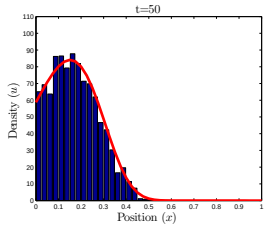
Comparison of the two models of average sensing on a domain with (linearly) density dependent growth

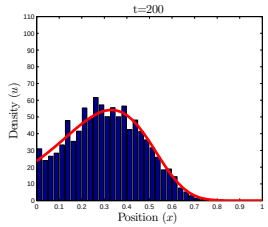
(Loading movie...)

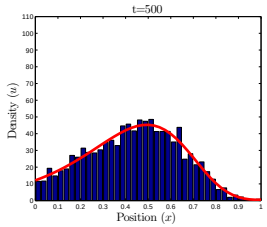
Histogram represents the average over 40 stochastic simulations and the red curve represnts the solution of the corresponding PDE.

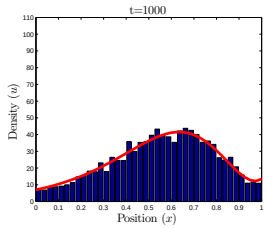
Summary

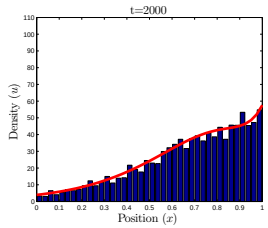
- Introduction to stochastic models for cell migration.
- Equivalence of stochastic and continuum models on a stationary domain.
- Different types of signal sensing.
- Equivalence of the two model types on growing domians.
- Different types domain growth.

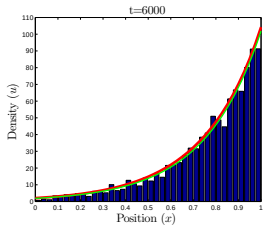












Further Work

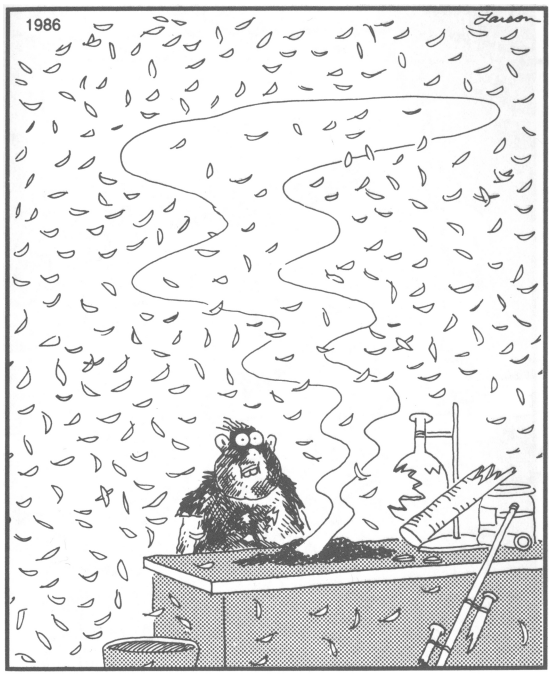
Where Next

- Acceleration of stochastic algorithms.
- Different boundary conditions i.e. non-zero flux.
- Cell death/birth*.
- Alternate signalling gradients*.
- Density dependent cell movement*.
- Extension to 2 and 3 dimensions.
- Different application areas e.g. brain/wound healing.

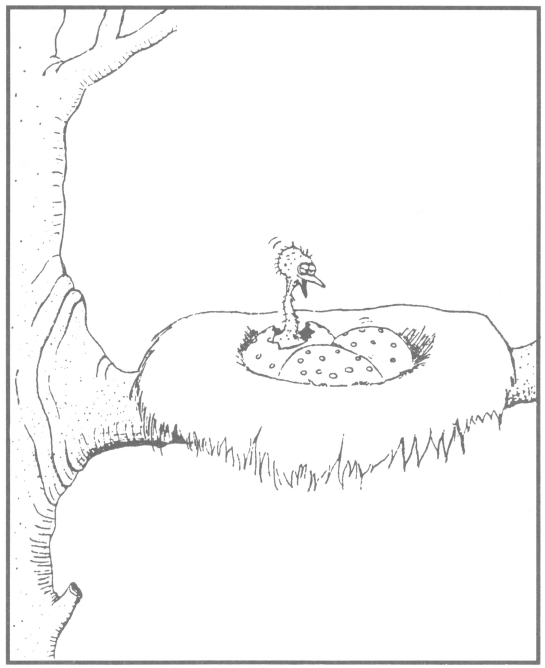
*These extensions have been implemented in R. E. Baker, C. A. Yates, and R. Erban. **From microscopic to macroscopic descriptions of cell migration on growing domains**.*Bull Math Biol*, Submitted.

Acknowledgements

- Dr Ruth Baker
- Dr Radek Erban
- Professor Philip Maini
- The Doctoral Training Centre (Oxford)
- EPSRC/BBSRC



God as a kid tries to make a chicken in his room.



"Last one out's a rotten egg!"

- And Gary Larson for these.