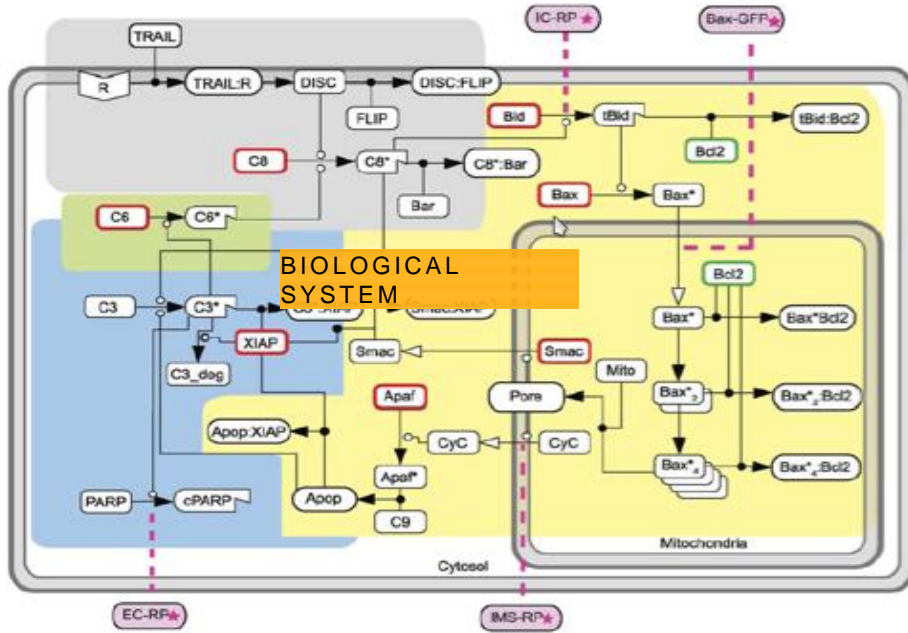


Why and How to use DBNs for biopathways?

Blaise Genest, SUMO

TRAIL-induced apoptosis: **complex system**

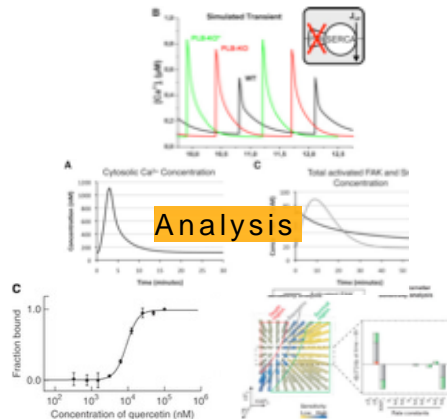
The « usual » Method



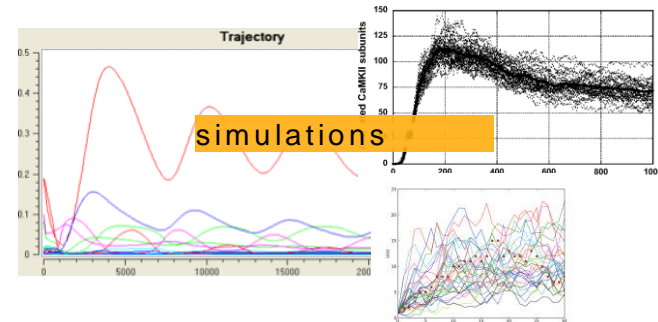
$$\begin{aligned} \frac{dc_1}{dt} &= -k_1 \cdot c_1 \cdot c_2 + k_2 \cdot c_3 \\ \frac{dc_2}{dt} &= -k_1 \cdot c_1 \cdot c_2 + k_2 \cdot c_3 + k_{17} \cdot c_{18} + k_{11} \cdot c_{11} \\ \frac{dc_3}{dt} &= k_1 \cdot c_1 \cdot c_2 - k_2 \cdot c_3 - k_3 \cdot c_3 \cdot c_4 + k_4 \cdot c_5 \\ \frac{dc_4}{dt} &= -k_3 \cdot c_3 \cdot c_4 + k_4 \cdot c_5 + k_{11} \cdot c_{11} + k_{20} \cdot c_{21} \\ \frac{dc_5}{dt} &= k_5 \cdot c_5 \cdot c_6 - k_6 \cdot c_7 - k_7 \cdot c_7 \cdot c_8 + k_8 \cdot c_9 \\ \frac{dc_6}{dt} &= -k_5 \cdot c_5 \cdot c_6 + k_6 \cdot c_7 + k_{11} \cdot c_{11} + k_{20} \cdot c_{21} \\ \frac{dc_7}{dt} &= k_5 \cdot c_5 \cdot c_6 - k_6 \cdot c_7 - k_7 \cdot c_7 \cdot c_8 + k_8 \cdot c_9 \\ \frac{dc_8}{dt} &= -k_7 \cdot c_7 \cdot c_8 + k_8 \cdot c_9 + k_{11} \cdot c_{11} + k_{20} \cdot c_{21} \\ \frac{dc_9}{dt} &= k_7 \cdot c_7 \cdot c_8 - k_8 \cdot c_9 - k_9 \cdot c_9 \cdot c_{10} + k_{10} \cdot c_{11} \\ &\quad - k_{15} \cdot c_9 \cdot c_{17} + k_{16} \cdot c_{18} \end{aligned}$$

MATHEMATICAL FORMALISM

ODE or stochastic equations



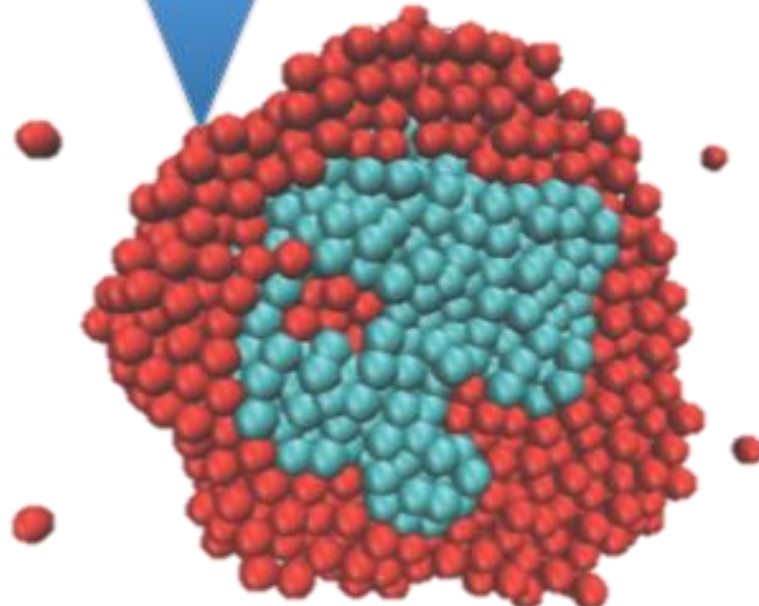
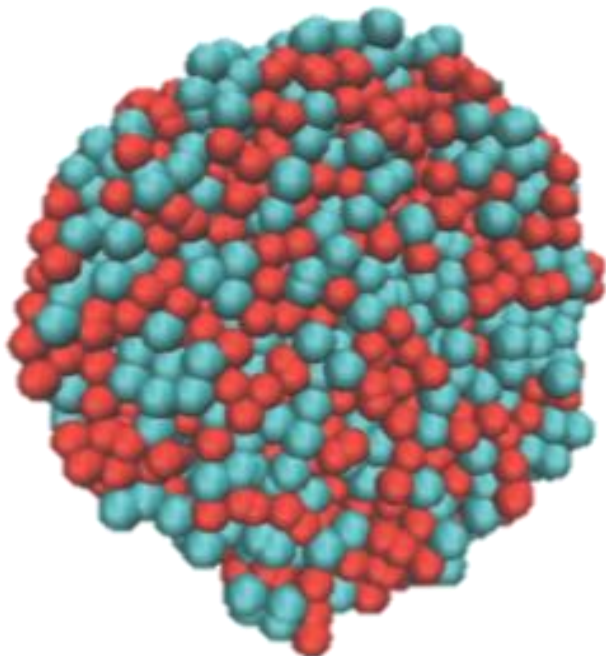
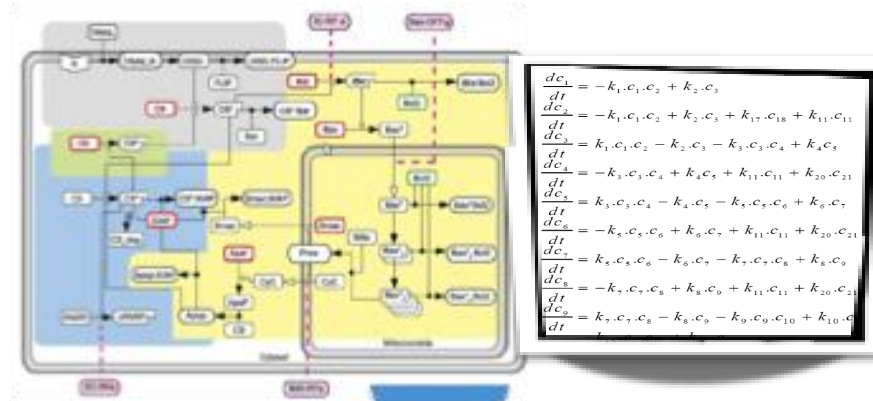
Analysis



simulations

TRAIL-induced apoptosis in spheroids

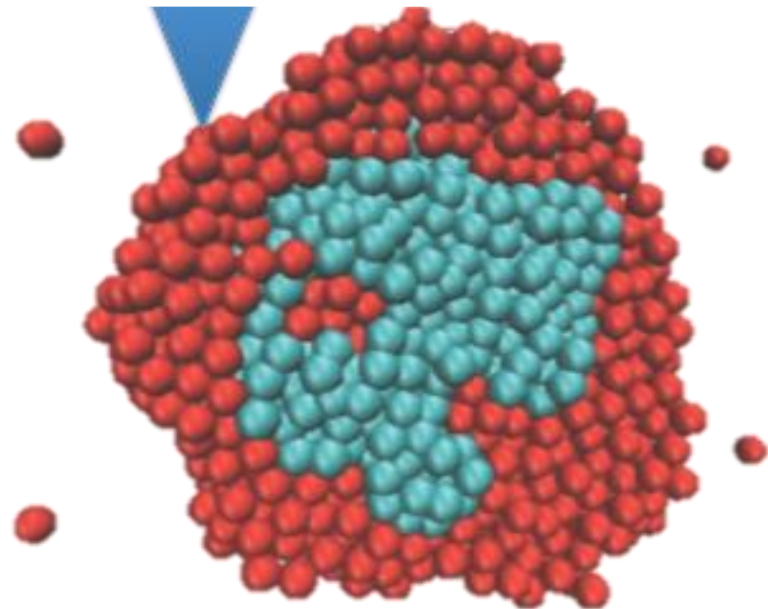
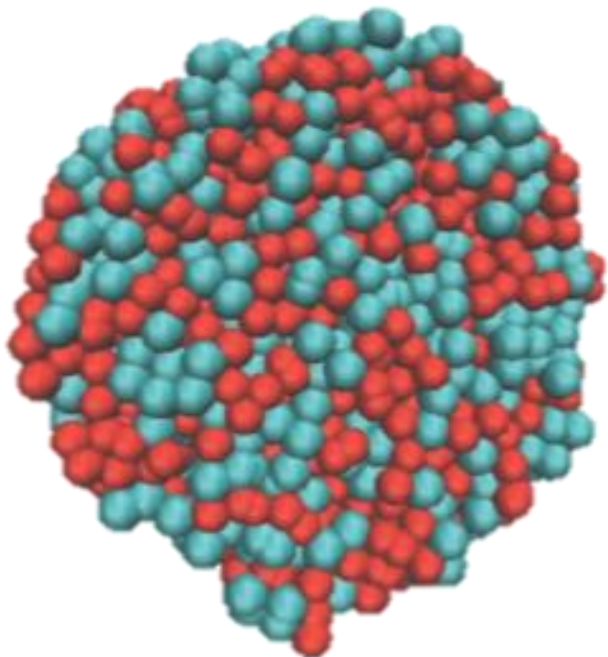
Multiscale problem: modeling and simulation is not efficient



TRAIL-induced apoptosis in spheroids

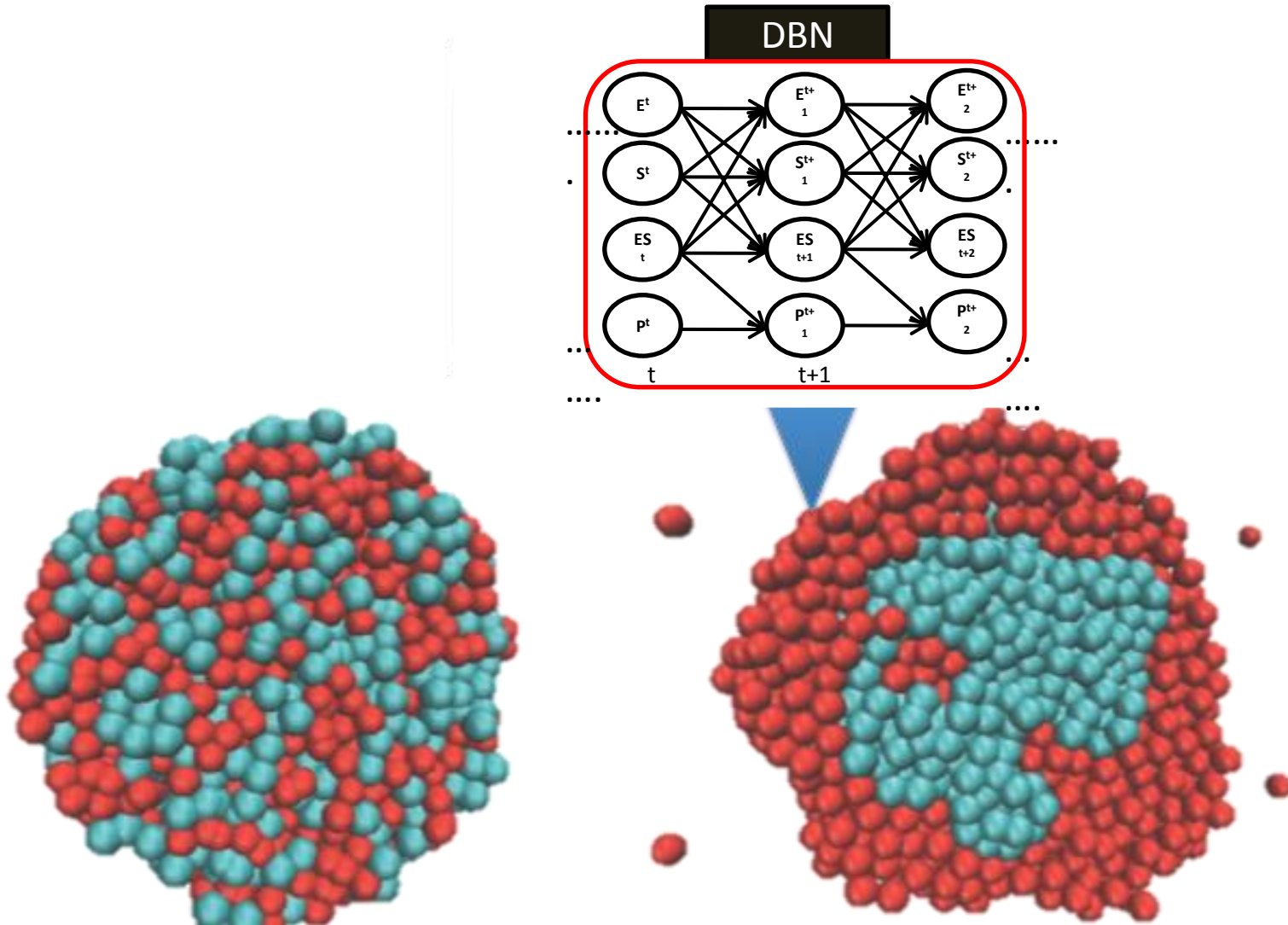
Multiscale problem: modeling and simulation is not efficient

(Approximate)
abstraction
of the low level
biochemical model



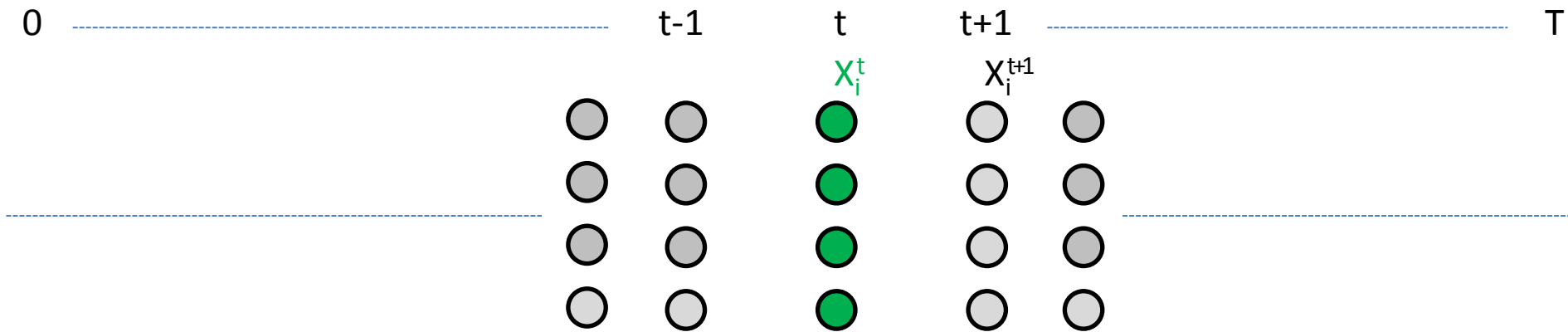
TRAIL-induced apoptosis in spheroids

Multiscale problem: modeling and simulation is not efficient



DBN = Approximate abstraction of the low level (biochemical) model

Dynamic Bayesian Networks (DBN)



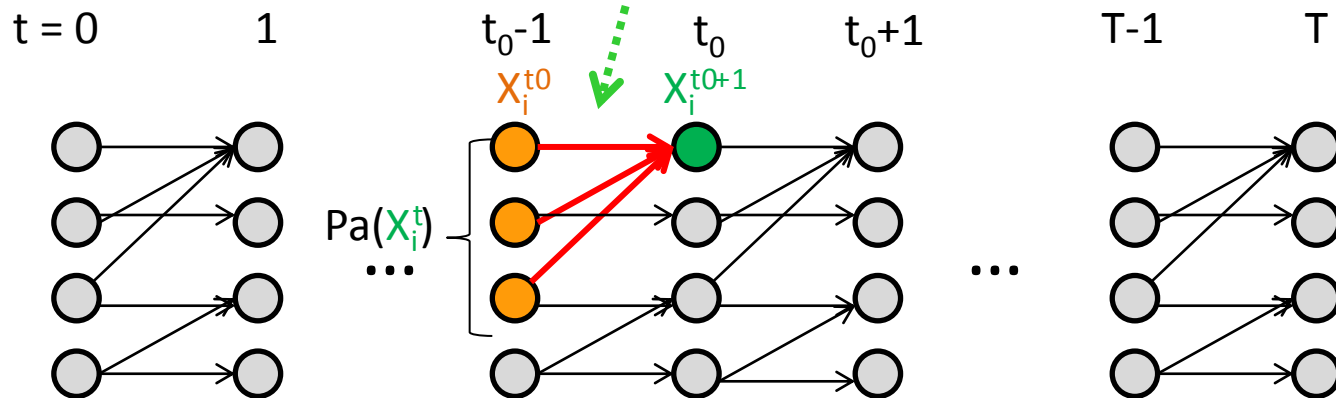
Discrete time: Finite number of time points (1 every 15 min).

Set of Nodes: each Node correspond to a biological species.
Dont need to have all species.

Each node associated with a small set of values
(« low/medium/high concentration »)

Conditional Probability Table (CPT)

| $X_1^{t_0}$ | $X_2^{t_0}$ | $X_3^{t_0}$ | $P(X_1^{t_0+1} = \cdot \mid X_1^{t_0}, X_2^{t_0}, X_3^{t_0})$ | | |
|-------------|-------------|-------------|---|------|------|
| | | | low | med | high |
| low | low | low | 0.1 | 0 | 0.9 |
| low | low | med | 0.1 | 0.4 | 0.5 |
| low | low | high | 0.7 | 0.1 | 0.2 |
| low | med | low | 0.2 | 0.02 | 0.78 |
| ... | | | | | |



ODE vs DBN

(stochastic) ODE

Continuous time
(simulation => every **ms**).

Biophysical model
(**All species**)

Continuous values
(concentration...)

1 trajectory

DBN

Discrete time
(1 time point every **15 min**).

Dont need to have all species.

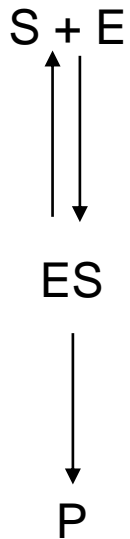
small set of values
(« low/medium/high concentration »)

Distribution/ Population of trajectory

DBN: high level abstraction

How to build a DBN?

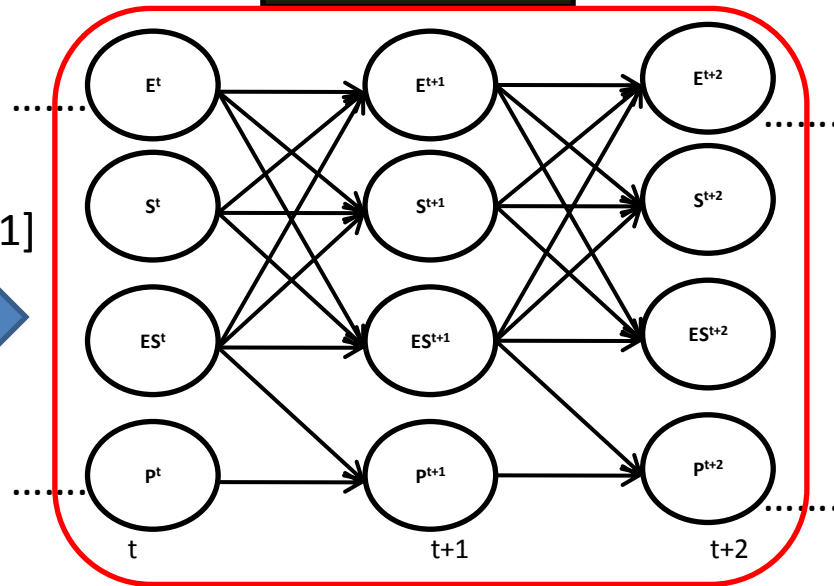
Biochem. equations



Liu.et.al [TCS 2011]



DBN



ODE simulations
+ statistics

Talk of Suche

| X_1^{t0} | X_2^{t0} | X_3^{t0} | $P(X_1^{t0+1} = \cdot \mid X_1^{t0}, X_2^{t0}, X_3^{t0})$ | | |
|------------|------------|------------|---|------|------|
| | | | low | med | high |
| low | low | low | 0.1 | 0 | 0.9 |
| low | low | med | 0.1 | 0.4 | 0.5 |
| low | low | high | 0.7 | 0.1 | 0.2 |
| low | med | low | 0.2 | 0.02 | 0.78 |
| ... | | | | | |

Semantic of DBN = Large Markov Chain

- We can use the CPTs to inductively compute the joint probability (a la Markov Chain):

$$P(\mathbf{X}^t = \mathbf{x}) = \sum_{\mathbf{u}} P(\mathbf{X}^{t-1} = \mathbf{u}) \left(\prod_i C_i^t(\mathbf{x}_i \mid \mathbf{u}_{\hat{i}}) \right)$$

CPT for node i at time t

- $\mathbf{u}, \mathbf{x} \in V^n \Rightarrow 5^{2n}$ computations to do.

\Rightarrow **Unfeasible**. Resort to approximated computations.

Talk of Matthieu

Statistics and Mutual Information

Analyzing precisely the ODE simulations in term of [mutual information](#)

-Which pairs of stochastic variables are more correlated?

=> Help building DBN ([Suchee](#)).

-Which subset of correlations to keep s.t. no big correlation is lost

=> Help the approximate computation to be accurate ([Matthieu](#))

Talk of Eric now