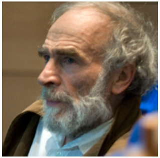
Tissue Transplantation Experiments: Inference and other Problems

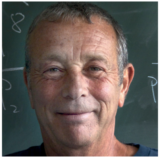
Yue Wang

Nov. 06, 2020

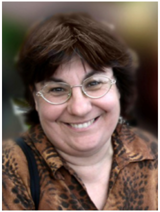
Mathematical biology group at IHÉS



Mikhail Gromov



Robert Penner



Nadya Morozova

Christophe Soulé

Yue Wang

Our recent works

- Yue Wang, Andrey Minarsky, Robert Penner, Christophe Soulé, and Nadya Morozova. (2020). “Model of morphogenesis.” *Journal of Computational Biology*, 27(9), 1373-1383.
- Theoretical models to explain early morphogenesis.

Our recent works

- Yue Wang, Jérémie Kropp, and Nadya Morozova. (2020). "Biological notion of positional information/value in morphogenesis theory." *International Journal of Developmental Biology*, in press.
- Logical analyses of some notions in developmental biology.

Our recent works

- Oksana Butuzova, Nikolay Pakudin, Yue Wang, Andrey Minarsky, Nikolay Bessonov, Robert Penner, and Nadya Morozova. (2020). “Formalization of embryogenesis as a developmental graph, and its application to phylogeny.” In preparation.
- Comparison between developmental processes. Define a metric on the space of trees, and design corresponding algorithms.

Our recent works

- Yue Wang, Jérémie Kropp, and Nadya Morozova. "Inference on tissue transplantation experiments." Preprint on arXiv: 2010.02704.
- Today's topic (simplified).

Experiments

- Tissue transplantation experiments.
- Take one piece of one tissue, and graft it to another tissue.
- Observe how the grafted tissue behaves.

Experiments

- To simplify the problem, we consider a rough classification of results:
- The grafted tissue develops normally.
- The grafted tissue develops abnormally.

Experiments

- Species: *Xenopus laevis* (African clawed frog).
- Donor: Upper lateral lip (development stage 11).
- Host: Lower lip (development stage 11).
- Result: Normal. The transplanted tissue will develop normally as if it was the host tissue.

Experiments

How to infer the unknown experimental results?

		Donor						
		AM19	PM19	PM15	UL11	LL11	LL15	LL19
	AM19	?	N	A	A	A	A	N
	PM19	?	N	?	N	N	?	?
H	PM15	?	?	?	?	?	?	?
o	UL11	?	N	?	N	N	?	?
s	LL11	?	N	?	N	N	?	?
t	LL15	?	?	?	?	?	?	?
	LL19	?	?	?	?	?	?	?

Table: Results reported by Krneta-Stankic et al. 2010

N=normal; A=abnormal. AM=anterior paraxial mesoderm; PM=presomitic mesoderm; UL=upper lateral lip; LL=lower lip; Number=developmental stage.

Ideas

- Core idea: similar experiments should have similar results.
- Similar experiments: similar donors and similar hosts.
- With this idea, we can use known results to infer unknown results.
- Assume we have known the similarities between experiments.

Ideas

Experiment similarity chart:
Black/red terms are experiments with known/unknown results.

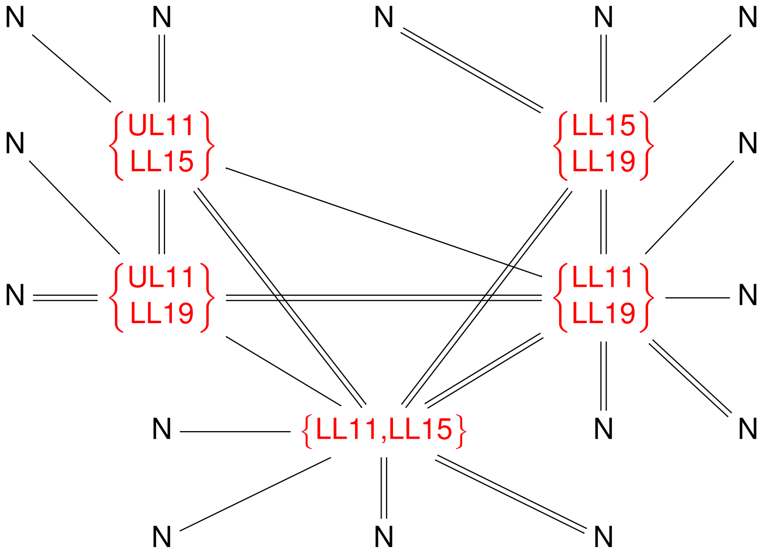
$\{\text{PM19, PM19}\} = \text{N}$ $\qquad\qquad$ $\{\text{PM15, PM15}\} = \text{N}$

$$\left\{\begin{matrix}\text{PM19}\\ \text{PM15}\end{matrix}\right\}$$

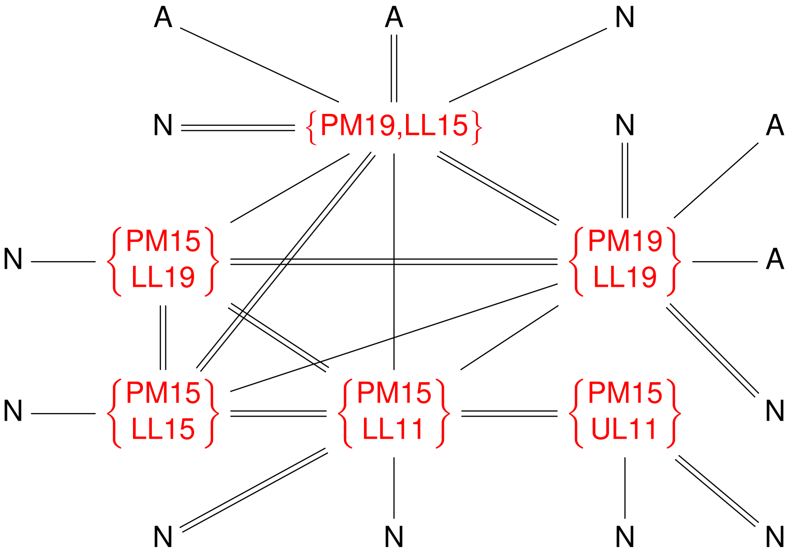
$\{\text{AM19, PM15}\} = \text{A}$ $\qquad\qquad$ $\{\text{AM19, PM19}\} = \text{N}$

Double/single/no line: high/medium/low similarity.

Ideas


N
N
N
N
N
UL11
LL15
LL15
LL19
N
N
N
UL11
LL19
LL11
LL19
N
N
{LL11,LL15}
N
N
N
N
N

Ideas

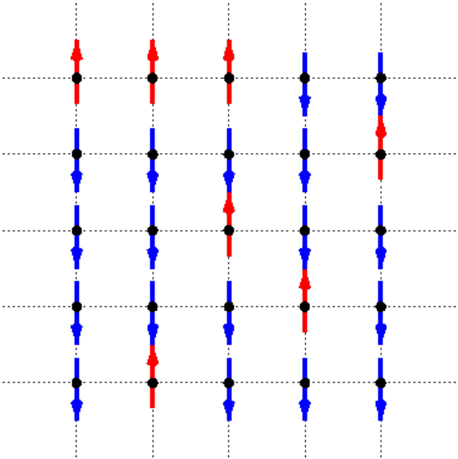

A
A
N
N
{PM19,LL15}
N
A
N
{PM15 LL19}
{PM19 LL19}
A
N
{PM15 LL15}
{PM15 LL11}
{PM15 UL11}
N
N
N
N
N

Ideas

- We design a penalty function that evaluates guesses of unknown experimental results.
- There is a penalty if two similar experiments have different results.
- For the concrete form of this function, we can get inspirations from the Ising model.

Ising model

The Ising model describes ferromagnetism in statistical mechanics. Consider a set of lattice sites, where each site k has a variable σ_k that takes $+1$ or -1.



Ising model

- For a configuration σ, its energy function is

$$\mathrm{H}(\sigma) = -\sum_{i \sim j} J_{ij}\sigma_i\sigma_j,$$

  where $i \sim j$ means site i and site j are neighboring, and J_{ij} is the interaction coefficient (no external field).
- The probability of a configuration σ is

$$P_\beta(\sigma) = e^{-\beta \mathrm{H}(\sigma)}/Z_\beta,$$

  where $\beta = (k_B T)^{-1}$, Z_β is the normalization constant.
- Configuration with high energy (high penalty) has small probability. Neighboring sites tend to have the same value.

Ideas

Analogies between tissue transplantation experiments and the Ising model:

Tissue transplantation	Ising model
Experiment similarity chart	Lattice
Experiment	Site
Result: normal/abnormal	Value: +1/-1
Similar experiments	Neighboring sites
Penalty: similar experiments have different results	Penalty: neighboring sites have different values
Penalty function?	Energy function

Pure analogy, not physical correspondence.

Ideas

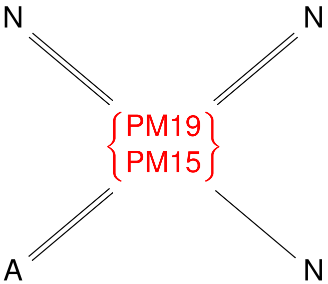
- Regard N as $+1$, and A as -1.
- If two experiments $\sigma_\gamma, \sigma_\delta$ are similar, set $J_{\gamma\delta} = 2J_0$ or J_0. Otherwise, set $J_{\gamma\delta} = 0$.
- The penalty function is

$$\mathrm{H}(\sigma) = -\sum_{\gamma,\delta} J_{\gamma\delta}\sigma_\gamma\sigma_\delta.$$

- The probability of a configuration σ is

$$P_\beta(\sigma) = e^{-\beta\mathrm{H}(\sigma)}/Z_\beta.$$

Ideas

- 

- $\{\text{PM19}, \text{PM15}\} = \text{N}$: $\mathrm{H} = -3$, $P = 0.65$.
- $\{\text{PM19}, \text{PM15}\} = \text{A}$: $\mathrm{H} = 3$, $P = 0.35$.
- Result=N is the most probable guess. $P(\text{N}) = 0.65$.

Ideas

Configuration					Penalty	Probability
-1	-1	-1	-1	-1	14	0.0019
1	-1	-1	-1	-1	10	0.0029
-1	1	-1	-1	-1	14	0.0019
1	1	-1	-1	-1	2	0.0064
-1	-1	1	-1	-1	14	0.0019
1	-1	1	-1	-1	2	0.0064
-1	1	1	-1	-1	6	0.0043
1	1	1	-1	-1	-14	0.0319
-1	-1	-1	1	-1	16	0.0016
1	-1	-1	1	-1	12	0.0024
-1	1	-1	1	-1	8	0.0035
1	1	-1	1	-1	-4	0.0117
-1	-1	1	1	-1	12	0.0024
1	-1	1	1	-1	0	0.0079
-1	1	1	1	-1	-4	0.0117
1	1	1	1	-1	-24	0.0868

Ideas

For each configuration of the unknown results (guesses), we can calculate its probability. We can determine the most probable guesses:

					Donor			
		AM19	PM19	PM15	UL11	LL11	LL15	LL19
	AM19	*N*	N	A	A	A	A	N
	PM19	*N*	N	N	N	N	N	N
H	PM15	*A*	N	*N*	N	N	N	N
o	UL11	*A*	N	N	N	N	N	N
s	LL11	*A*	N	N	N	N	N	N
t	LL15	*A*	N	N	N	N	*N*	N
	LL19	*N*	N	N	N	N	N	*N*

Ideas

We can also calculate the expectation of all guesses, i.e., the probability for each experimental result to be "N":

		Donor						
		AM19	PM19	PM15	UL11	LL11	LL15	LL19
	AM19	*100%*	100%	0%	0%	0%	0%	100%
	PM19	*100%*	100%	65%	100%	100%	49%	56%
H	PM15	*0%*	65%	*100%*	62%	62%	53%	54%
o	UL11	*0%*	100%	62%	100%	100%	81%	81%
s	LL11	*0%*	100%	62%	100%	100%	90%	90%
t	LL15	*0%*	49%	53%	81%	90%	*100%*	86%
	LL19	*100%*	56%	54%	81%	90%	86%	*100%*

Another situation

What if the known experimental results are not deterministic, but stochastic?

Another situation

		Donor				
		PLE11	PLE12	PLE14	PLE16	PLE19
	LFR\PLE14	61%	58%	82%	?	?
Host	LFR\PLE16	?	?	?	?	?
	LFR\PLE19	4%	24%	83%	?	100%

Table: Results reported by Henry et al. 1987

- Percentage is lens formation (normal growth) rate.
- PLE11: presumptive lens ectoderm, stage 11.
- LFR\PLE14: lens-forming region without presumptive lens ectoderm, stage 14.

Another situation

Decompose the stochastic results into several deterministic results with different probabilities:

$$[61\%\text{N } 58\%\text{N}] : P([\text{N N}]) = 61\% \times 58\% = 31\%.$$

$$P([\text{N A}]) = 61\% \times (100\% - 58\%) = 26\%.$$

$$P([\text{A N}]) = (100\% - 61\%) \times 58\% = 23\%.$$

$$P([\text{A A}]) = (100\% - 61\%) \times (100\% - 58\%) = 16\%.$$

$$[61\%\text{N } 58\%\text{N}] = 35\%[\text{N N}] + 26\%[\text{N A}] + 23\%[\text{A N}] + 16\%[\text{A A}].$$

For each deterministic configuration, apply our method to obtain the expectation of guesses. Then average over these deterministic configurations.

Another situation

		Donor				
		PLE11	PLE12	PLE14	PLE16	PLE19
	LFR\PLE14	61%	58%	82%	93%	94%
Host	LFR\PLE16	39%	53%	88%	97%	97%
	LFR\PLE19	4%	24%	83%	96%	100%

Experimental design

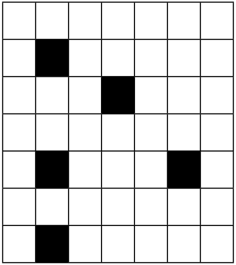
- There are many tissue transplantation experiments, and we want to know all the results.
- We can choose some experiments to conduct, and use the known results to infer other experiments.
- How to choose experiments to conduct?
- We need enough data to conduct the inference. We should minimize the experimental cost.

Experimental design

- The results of non-conducted experiments are inferred by similar conducted experiments.
- To guarantee the inference quality, one non-conducted experiment should be similar to at least k conducted experiments (e.g., $k = 2$ or $k = 1$).
- The most efficient design: no conducted experiments are similar, and each non-conducted experiment is similar to exactly k conducted experiments.

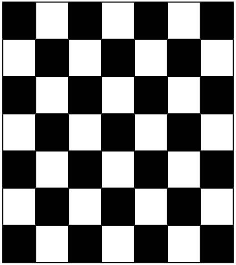
Experimental design

- Consider the figure that each unit is an experiment, and neighboring units are similar experiments.
- Black units are conducted experiments, and white units are non-conducted experiments.
- How to color the figure, so that two black units are not neighboring, and each white unit is neighboring to k black units?

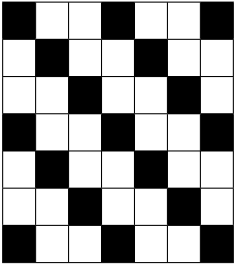


Experimental design

No neighboring black units, and each white unit is neighboring to $k = 4$ black units (ignore the boundary cases).

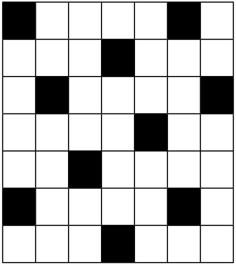


No neighboring black units, and each white unit is neighboring to $k = 2$ black units (ignore the boundary cases).



Experimental design

No neighboring black units, and each white unit is neighboring to $k = 1$ black units (ignore the boundary cases).



Experimental design

- For $k = 4$, we need to conduct $1/2$ experiments.
- For $k = 2$, we need to conduct $1/3$ experiments.
- For $k = 1$, we need to conduct $1/5$ experiments.

Experimental design

- In practice, the experiment similarity chart is not 2-dimensional, but 4-dimensional.
- Similar coloring problems for such 4-dimensional figures.
- We need some abstract methods.

Experimental design

- For $k = 8$, color a unit (x, y, z, w) if

$$x + y + z + w \equiv 0 \mod 2.$$

  We need to conduct $1/2$ experiments.

- For $k = 4$, color a unit (x, y, z, w) if

$$x + y + z + w \equiv 0 \mod 3.$$

  We need to conduct $1/3$ experiments.

Experimental design

- For $k = 2$, color a unit (x, y, z, w) if
$$x + 2y + z + 2w \equiv 0 \mod 5.$$
  We need to conduct $1/5$ experiments.
- For $k = 1$, color a unit (x, y, z, w) if
$$x + 2y + 3z + 4w \equiv 0 \mod 9.$$
  We need to conduct $1/9$ experiments.

Experimental design

- In practice, $k = 2$ or $k = 1$ is enough to conduct satisfactory inference. Therefore we only need to conduct $1/5 - 1/3$ experiments (two-dimensional) or $1/9 - 1/5$ experiments (four-dimensional).
- The more experiment similarities we have, the fewer experiments we need to conduct.

Thank you!