Bachelorthesis about Community Structure in Graphs

Using Network Propagation Methods and Applications in Protein Function Annotaions

Yiftach Kolb

Instructor

Prof. Dr. Matin Vingorn

September 9, 2020

1 Introduction

This Thesis deals with the propagation methods for clustering and community structue in Graphs. The original motivation is to find application of these methods on Protein-Protein Interaction Networks (PPI) specifically for function annotations.

We are given an (in my case undirected) connected graph, which represents some sort of network structure, be it a PPI or A social network. Real world networks tend to have some characteristics that are clearly not random, such as the 'small world property' and tendency to forming hubs.

2 Linear Algebra Primer

Definition 2.1. A **Transition Matrix** is a real valued non negative square (n^2) matrix A s.t. each of its columns sums up to one:

$$\forall j \sum_i A_{i,j} = 1,\ \forall i,j A_{i,j} \geq 1$$

A acts from the left as a linear mapping: $T(v) := Av, \forall v \in R^n$. In the following script we might interchangeably and indistinguishably use T (the mapping) or A the matrix.

A transition is **positive**, designated by $A > 0$ if all its entries are positive.

A transition is **regular** if for some $k > A^k$ is positive. The same property is called **primitive** in some other sources.

A transition is **irreducible** if for every entry $A_{i,j}$ there is some k such that $A^k_{i,j} > 0$ (More on that further down).

Equivalently it can be shown that A matrix A is irreducible by if and only if it is NOT similar by permutations matrices to a block upper triangular matrix which means $\nexists P : A = P \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} P^{-1}$

Where P is a permutation matrix and B, C are square matrices of size greater than 0.

Definition 2.2. A **State** is a non-negative vector $v \in R^n$ s.t $\sum_i v_i = 1$.

Remark 1. If v is a state and A is a transition as defined in 2.1, then Av is also a state because:

$$\sum_i (Av)_i = \sum_j v_j (\sum_k A_{j,k}) = \sum_j v_j \cdot 1 = 1$$

If A is a transition and x, y are two states such that $x \leq Ay$ then $x = y$.

Definition 2.3. Given $A \in \mathbb{C}^{n \times n}$ We let $|A| \in \mathbb{C}^{n \times n}$ be the resulting matrix from applying $|\cdot|$ element wise. Given a vector $v \in \mathbb{C}^n$ we let $|v| \in \mathbb{C}^n$ the corresponding non-negative vector.

We also let $A > 0, v > 0$ mean that it holds coordinate wise.

Here is a very useful lemma for non-negative matrices which we will need later:

Lemma 2.1. *Let* $0 \leq A \leq B \in \mathbb{R}^{n \times n}$ *and let* $0 < v \in \mathbb{R}^n$. *If* $Av = Bv$ *then* $A = B$.

Proof. Trivial. □

Remark 2. If $u \in \mathbb{C}^n$ is on the unit circle and T a transition then $|u|$ is a transition, meaning $\||u|\| = 1$, so $T|u|$ is also a transition so $\|T|u|\| = 1$.

We have (component-wise) $|Tu| \leq T|u|$. If $T > 0$ and u has negative or non-real entries, then this inequality must be strict and then $\|Tu\| < \|T|u|\| = 1$.

Lemma 2.2. *If T is a transition, then there is a state u such that $Tu = u$.*

Proof. Because the columns of A all sum to 1, the columns of $A - I$ all sum to 0. Therefore $(1, 1, \ldots, 1)$ is an (left) eigenvector of the rows with eigenvalue 1. Therefore there is also some real (right) column eigenvector with eigenvalue 1. (Also it follows from the Brouer fixed point theorem because T maps the l_1 sphere to itself).

Let $u \in R^n$ be such vector: $Au = u$. Let $u = u^+ + u^-$ such that $u^- = \min(u_i, 0)$ and the other one defined similarly for the non-negative components.

Because A is non-negative $A(u^+) \geq (Au)^+ \geq 0$, and $A(u^-) \leq (Au)^- \leq 0$.

From A being non-negative and $(Au)^+ + (Au)^- = Au = u = u^+ + u^-$ And also $(Au)^+ = (u)^+$, so we must have $Au^+ \geq (Au)^+ = u^+$ (component wise). But if we had a strict inequality we would get: $\|A(u^+/\|u^+\|_1)\|_1 > 1$ which is a contradiction to A being a transition matrix.

Then $Au^+ = u^+, Au^- = u^-$ and one of them must be non-zero. It follows that A has a non-negative eigenvector with eigenvalue 1 (one of $u^+, -u^-$ which is non-zero). If we l_1-normalize that eigenvector it becomes a state. □

Lemma 2.3. *If a transition $A > 0$ (or primitive) , then it has exactly one eigenvector v with eigenvalue 1 and in addition it can be chosen so that $v > 0$ Any for any other eigenvalue λ it holds that $|\lambda| < 1$.*

If A is just irreducible then then again $v > 0$ and is unique but there may be additional eigenvalues on the unit circle.

Proof. Let $A > 0$ be a transition. We already know that there exists at least one such eigenvector. Let u, v s.t $Au = u, Av = v$. We can assume these are real vectors because A has only real entries. Therefore we can choose u, v to be states as we have already proven above.

Then let $w = u - v$. So $Aw = A(u - v) = u - v = w$. And $\sum_i w_i = 1 - 1 = 0$ by choice of u, v.

Like before $Aw^+ = w^+, Aw^- = w^-$ and because $w \neq 0$ but sums up to 0, both $w^+, -w^- > 0$. Because w^- is non zero exactly on entries where w^+ is zero and vice versa, each of them must have at least one 0 entry (and one none zero). But because A is strictly positive and w^+ is non-negative, Aw^+ must have ONLY positive entries, contradicting $Aw^+ = w^+$. It follows then that $u - v = 0$ is the only possibility, hence the eigenvector with 1 as eigenvalue is unique.

Suppose there is $Aw = \lambda w$ where $|\lambda| = 1$. Choose w so it is l_1-normalized. Then $|w| = |Aw| \leq A \cdot |w|$ If w has any negative or complex coordinantes, then $|Aw| < A|w|$ and therefore $1 = \||Aw|\|_1 < \|A|w|\|_1 = 1$, a contradiction. Therefore there cannot be any other eigenvalues on the unit circle.

Extending this for primitive matrix is easy because for some sufficiently big k $A^k > 0$.

To prove the uniqueness of the 1-eigenvector for the irreducible case, we have $(\forall k \in \mathbb{N})A^k w^+ = w^+$ and from that with some more work left undone it follows that $w^+ > 0$ or $w^+ = 0$. □

□

Remark 3. The lemmas and theorems in this section are phrased in term of transitions. They hold true almost verbatim in the more general case of positive/non-negative linear transformations.

In general a non-negative linear transformation has a **spectral radius** $\rho = \rho(A)$ which is the absolute value of its greatest eigenvalue. In the case of positive maps there is a unique single eigenvector with ρ as the unique greatest eigenvalue and so forth .... When we deal with a transition map, lemma 2.3 guaranties it has a spectral value of $\rho(A) = 1$.

Theorem 2.1. *Let T be a positive (or primitive) transition. Then*

1. 1 is the greatest eigenvalue of T and it has one unique eigenvector which is also positive, so there exists a unique stationary state.

2. All the other eigenvalues have absolute value strictly less than 1.

3. For every state u, $T^k u$ converges to the stationary state π. In particular the columns of T^k converge to π.

Proof. 1 and 2. We already know.

3. There is a Jordan decomposition $T = PJP^{-1}$.Such that J is a Jordan matrix, $j_{1,1} = 1$ and the rest of the main diagonal $|J_{i,i}| < 1$. So now the convergence is clear $J^k \to (e_1|0\ldots|0)$. For the matrix P it must hold that $P = (P_1, \ldots, P_n)$ and P_1 is the eigenvector of T corresponding to 1 which we are free to l_1 normalize and the first row of P^{-1} is the left eigenvector corresponding to 1 and so force ....

Some more work or literature check should confirm that $T^k \to (v|v\ldots|v) = V$. Then one can verify $Vu = v$ for any state u. □

Theorem 2.2. *Let T be an irreducible positive (or primitive) transition. Then:*

1. Then 1 is the greatest eigenvalue of T and it has one unique eigenvector which is also positive, so there exists a unique stationary state.

2. If there are other other eigenvalues on the unit circle then their algebaic multiplicity is equal their geometric multiplicity.

*3. For every state u, the **Cesaro sums** $\frac{1}{n}\sum_{k=1}^{n} T^k u$ converge to the stationary state π.*

Proof. 1 We already know.

2. There is a Jordan decomposition $T = PJP^{-1}$. such that J is a Jordan matrix, $j_{1,1} = 1$ and the rest of the main diagonal $|J_{i,i}| \leq 1$.

If we had a Jordan block of size greater than 1 for an eigenvalue λ, Then on the superdiagonal of J^k we would have $k\lambda$. If $|\lambda| = 1$ then J^k and hence T^k would be unbounded, but that is impossible since T^k is a transition. If follows that all eigenvalues on the unit circle must be simple (alebgaic multiplicity equals geometric).

3. For the convergence, it follows from calculations on the Jordan blocks, which I omit. See Meyer [3] or Serre [5] for rigorous proof. □

What differs irreducible non-primitive matrices from being primitiv is that they are periodical on their complex eigenvectors with complex eigenvalues on the unit cycle. There is, in fact a wonderful theorem from Wielandt which characterizes these Matrices:

Theorem 2.3 (Wielandt (1950))**.** *Let $A, B \in \mathbb{C}^{n\times n}$ such that $A \geq 0$ is irreducible and $|B| \leq A$ (component-wise). Then $\rho(B) \leq \rho(A)$. If $\rho(B) = \rho(A)$ then $|B| = A$, B has an eigenvalue of the form $\mu = \rho(A)e^{i\phi}$ and:*

$$B = e^{i\phi} DAD^{-1}$$

where D has the form:

$$D = \begin{pmatrix} e^{\theta_1} & & & \\ & e^{\theta_2} & & \\ & & \ddots & \\ & & & e^{\theta_2} \end{pmatrix} \tag{1}$$

And conversely any B of the form 1 has $\rho(B) = \rho(A)$, $|B| = A$ and μ is an eigenvalue of B which corresponds to the eigenvalue $\rho(A) = |\mu|$ the greatest eigenvalue of A.

Proof. To see a rigorous proof I suggest Meyer [3].

The keys for proving this theorem are: First WLG assume A is a transition. This is possible because we may replace A with $W^{-1}A$, and B with $W^{-1}B$, where W is the diagonal matrix that has the column-sums of A. Since A is irreducible it cannot have a column or a row that is all 0 so this diagonal is positive W is indeed invertible and later we can cancel out the W's and return to the general case.

Let v be the μ-eigenvector $Bv = \mu v, \|\mu| = 1$ and choose it so that $\|v\|_1 = 1$.

Then

$$|v| = |\mu v| = |Bv| \leq |B||v| \leq A|v| \tag{2}$$

Since A is a transition by remark 1 $A|v| = |v|$. Since A is irreducible and $A|v| = |v|$ we must have by 2.2 that $|v| > 0$, and then by lemma 2.1 $A = |B|$ so that proves the first part.

Now let $w = v/|v|$ (component-wise division) and let

$$D = \operatorname{diag}(w) = \begin{pmatrix} e^{\theta_1} & & & \\ & e^{\theta_2} & & \\ & & \ddots & \\ & & & e^{\theta_2} \end{pmatrix}$$

Then $v = D|v|$, $|v| = D^{-1}v$ and we have:

$$\begin{aligned} A|v| = |v| = D^{-1}v = \\ = D^{-1}\mu^{-1}Bv = \mu^{-1}D^{-1}BD|v| \end{aligned} \tag{3}$$

We know already that $|v| > 0$. If $C := \mu^{-1}D^{-1}BD$ contains any negative or complex entries, then 3 cannot hold. It follows that

$$A = C = \mu^{-1}D^{-1}BD$$

This proves the harder direction of the claim, the other direction is easy □.

□

The amazing consequence from 2.3:

Theorem 2.4 (Corollarly)**.** *If A is irreducible transition with h eigenvalues on the unit circle then its eigenvalues are exactly the hth unit roots $\lambda_k = e^{2\pi ik/h}, k = 0 \dots n-1$ and A is similar to λA by a diagonal matrix for any such eigenvalue.*

Proof. Use theorem 2.3 with $B = A$. If $|\lambda| = 1$ is an eigenvalue then $A = \lambda DAD^{-1}$. Since similarity doesn't change the eigenvalues, A and λA must have the same eigenvalues with the same multiplicity. Since 1 is a simple eigenvalue of A and hence of λA, and its corresponding eigenvalue λ is simple in λA and therefore in A as well.

The matrices $\lambda_k A$, $k = 0 \dots h$ are all similar, with $\lambda_0 = 1$ and $|\lambda_k| = 1, k = 0\dot{h} - 1$ all simple eigenvalues. The only way for this to hold is if $\{\lambda_k\}_0^{h-1}$ form a multiplicative group of order h on the unit circle, in other words, the eigenvalues are exactly all the hth unit roots. □

Remark 4. If A is irreducible transition with exactly h eigenvalues on the unit circle then h is called the **period** of A. $A \geq 0$ is primitive if and only if it is irreducible and aperiodic ($h = 1$).

If $\omega = e^{2\pi i/h}$ then 2.4 shows that $A = \omega DAD^{-1}$. So if λ is any eigenvalue not necessarily on the unit circle, then $\omega\lambda$ is also an eigenvalue with the same multiplicity and rotation with ω is an automorphism on the eigenvalues.

We may choose D so that $D[1,1] = 1$. If we reindex the dimension we can make D look like

$$D = [D_0|\omega D_1| \dots \omega^{h-1} D_{h-1}]$$

So Indexes corresponding to the same phase of the period appear sequentially.

Then use the identify $\forall k D^h A^k = A^k D^h$ and the fact that A is irreducible to show that $D^h = I$. Then use the identity $\omega DA = AD$ to show that in the new indexing A has the following periodical block structure (0 on the main diagonal):

$$A = \begin{pmatrix} 0 & M_1 & 0 & \dots & 0 \\ 0 & 0 & M_2 & 0 \dots & 0 \\ & & \ddots & \ddots & \\ 0 & \dots & 0 & 0 & M_{h-1} \\ M_h & 0 & \dots & 0 & 0 \end{pmatrix}$$

Now we will just present the Perron-Frobenius theorems. The main parts that are important to our work have appeared in the previous theorems.

Theorem 2.5 (Perron-Frobenius [3])**.** *Let $0 < A \in \mathbb{R}^{n \times n}$ with spectral radius $\rho := \rho(A)$, then the following are all true:*

- $\rho > 0$
- *ρ is a simple root of the characteristic polynomial of A, in other words its algebraic multiplicity is* 1.
- $(\exists v > 0) Av = \rho v$
- *If $Au = \lambda u$ and $\|u\| = \rho$ then $\lambda = \rho$ namely, ρ is the unique eigenvalue on the spectral circle.*
- *(Collatz-Wielandt Formula) $\rho = \max_{x \in \Gamma} \min_{i:x_i \neq 0} [Ax]_i / x_i$ with $\Gamma = \{x | x \geq 0, x \neq 0\}$*

Theorem 2.6 (Perron-Frobenius for irreducible matrices [3]). *Let $0 \leq A \in \mathbb{R}^{n\times n}$ be irreducible with spectral radius $\rho := \rho(A)$, then the following are all true:*

- $\rho > 0$
- *ρ is a simple root of the characteristic polynomial of A, in other words its algebraic multiplicity is* 1.
- $(\exists v > 0)Av = \rho v$
- *There are no additional non-negative unit eigenvectors of A other than v.*
- *(Collatz-Wielandt Formula)* $\rho = \max_{x\in\Gamma} \min_{i:x_i\neq 0}[Ax]_i/x_i$ *with* $\Gamma = \{x|x \geq 0, x \neq 0\}$

3 More on Matrices, Graphs and Stochastics

A directed non-weighted graph G can be uniquely represented by its **adjacency matrix**, $A_{i,j} := 1$ if and only if there is a directed edge from i to j (if we want to use it for transitioning on columns as done above). It's possible to assign different edge weights rather than only 1 and 0. If the graph is undirected each edge would be counted in both directions and the matrix is symmetric. Relabeling the vertices of a graph yields an adjacency matrix that is similar by permutations (PAP^{-1}, where P is a permutation matrix) and vice versa.

To turn A into a transition, normalize each column by dividing it with its out going rank, so let $D_{i,i} = \text{outrank}(i)^{-1}$, AD is the transition matrix of this graph (because right-multiplying by D normalizes each column by its rank). If the graph was stochastic to begin with then the adjacency matrix as we defined it is already column-normalized. Going forward when we are always implicitly assuming that the adjacency matrix is column-normalized.

Definition 3.1. A graph is G **strongly connected** if there is a directed path form any edge to any edge. if and only if the adjacency matrix is irreducible.

We say that the **Period of a vertex** $v \in V(G)$ is the greatest common divider of all closed paths on v. If the periods of all vertices are equal (spoiler- in the case that G is strongly connected they do), we call it the **Period of the graph** G.

Remark 5. If G is strongly connected then the periods of all vertices are indeed equal and its easy to prove. The corresponding adjacency matrix A is irreducible so it too has a period h as defined in 4 and it is equal to the graph period (can be shown using 2.4 and 4).

So if the graph G is strongly connected and has period 1 then the adjacency matrix is aperiodical and hence primitive, and vice versa.

From all the above we have seen that a Markov process can be represented in two equivalent ways —as a transition matirx and as the the corresponding weighted directed graph.

If the graph G is strongly connected and aperiodic, its corresponding adjacency matrix is primitive. We know from 2.5 that there is a unique stationary distribution p and that the Markov converges to p no matter from which distribution it starts. We may calculate p using the **power method** which is efficient because it can be done in a matrix-free method. We don't need to know the matrix itself just the dot product of it with a state vector.

If the graph G is strongly connected, then 2.6 assures as the existence and uniqueness of a stationary distribution p. But if G is not aperiodic, the corresponding adjacency matrix is not

primitive. We cannot use the efficient power method to calculate p. Also the process itself doesn't stabilize on p, it is periodic and cycles on between the h eigenvectors on the unit circle.

We are therefore interested to find next how to convert an imprimitive matrix (= irreducible but not primitive) to a primitive matrix or or equivalently a strongly connected but periodic graph to an aperiodic graph.

Lemma 3.1. *Let $A \geq 0$ be irreducible. Then $(A+I)^{n-1} > 0$, and therefore $A+I$ is primitive.*

Proof. Notice that the I represents self loops and it absorbs all the lower powers so if $(A^k)_{i,j} > 0$ for some $k < n$ then so is $(A+I)^{n-11}$. Let $G := G_A$ be the corresponding graph to A. Then G is strongly connected. For every $i \neq j$ there is a directed **shortest path** σ in G from i to j. Its length must be $|\sigma| < n$. Otherwise σ would have to contain a cycle and not be of minimal length. This shows that we have for every $i \neq j$ some k sucht that $A^k_{i,j} > 0$ and therefore $(A+I)^{n-1} > 0$. □

The properties of irreducibility, primitivity and positivity only depend on the sign $(-, 0, +)$ of the entries and not on their size. So we use the following definition to test matrices for these properties:

Definition 3.2. Let $A \in \mathbb{R}^{n\times n}$, then its binary form is the matrix $\beta(A) := \operatorname{sgn}(A)$ where sgn is applied element-wise.

The following trivial lemmas would help as to construct primitive transitions:

Lemma 3.2. *A is positive/primitive/irreducibly if and only if $\beta(A)$ is.*

Let $0 \leq \beta(A) \leq \beta(B)$. Then If A is positive/primitive/irreducibly so is B.

Let $0 < \alpha < 1$. If T, S are transitions the so is $W = (1-\alpha)T + \alpha S$. If one of S, T is positive/primitive/irreducible so is W

Let G be any weighted graph and A its adjacency transition matrix. Some vertices may be unreachable from other vertices and we there might not exist a single and unique stationary distribution. A random walk on this graph is generally dependent on the initial state. However if we allow the possibility of 'random restart' from any state, this graph becomes totally connected it is guarantied to have a unique stationary distribution to which any random walk converges regardless of the initial state.

In terms of matrices, we create a positive transition matrix $W = (1-\alpha)A + \alpha\mathbf{1} > 0$. The stationary distribution is called the **PageRnak** for G with restart parameter α. It is used to order the vertices according to their 'relevance' in the network.

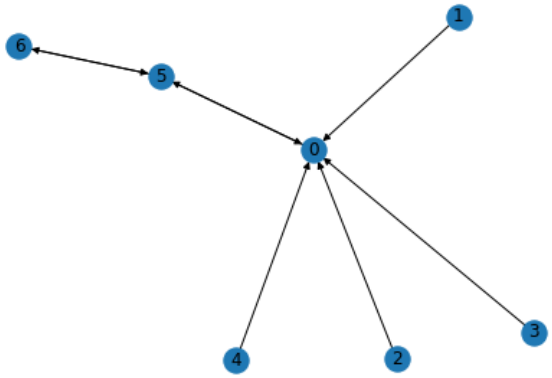

Figure 1: This is an example of a weakly connected graph. If we start walking from 0, 5 or 6 we can never reach vertices $1-4$. It is also not immediately clear which vertex 0 or 5 is more 'important'. While 0 is directly connected to more vertices, 5 may get more 'flow' through it since every path of length 2 or more passes though it.

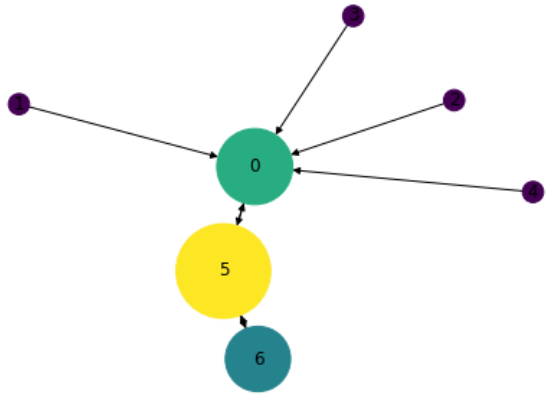

Figure 2: The same graph with vertex size and color indicating its PageRank with parameter $\alpha = 0.15$. We see that 5 is ranked first, followed by 0. The smaller the restart parameter, the more important 5 will get because we allow for longer paths with fewer restarts.

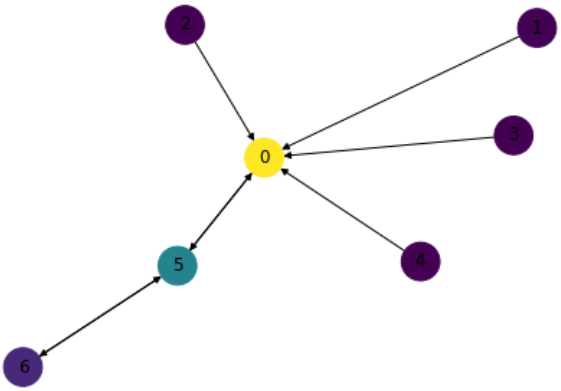

Figure 3: When the restart parameter is too large, the rank becomes almost uniform because the convex combination of the original graph with the n-clique graph of the restarts is weighted too heavily on the latter

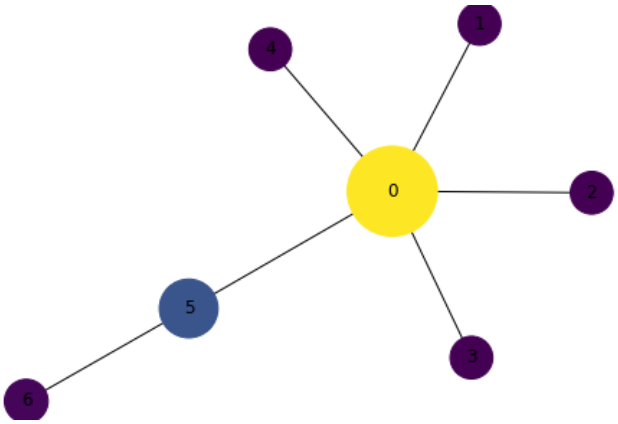

Figure 4: When we treat the graph as undirected and calculate the PageRanks ($\alpha = 0.15$), Vertex 0 comes this time on top. It is more central than 5 and more random paths intersect it than any other vertex.

4 The Propagation Method

The normalize adjacency matrix of a connected graph (for simplicity we use here undirected) is a transition matrix and represents a Markov process. On a given time-step, a visitor on a graph node chooses his next station at random among the neighbours of the current station (node).

We want to know first, if we repeat this process to infinity will the frequency of the visits at each node stabalize

5 testing

5.a citations

For more info see (textcite) Meyer [3, p. 631].

This is a normal (cite) [1, p. 115].

This is an (autocite) see [2, p. 231].

Change autocite in the usepackage definition to suit your needs. Use textcite if you want to insert inline cite that includes the author's name.

Here is a [4] cite from wikipedia.

Just cite is just the normal citation command and nocite will make sure that the reference appears in the bibliography even though it may have not specifically been mentioned in the text somewhere.

5.b Images

Here is an image

Refer to it with fig 5 tada.

5.c Tables

Table 1: Nernstsche Gleichgewichtspotenziale und das resultierende Membranpotenzial nach Goldmann Gleichung

Ion	rel' Permeabilität	Konz. in	Konz. auß	GG Potenzial
K^+	1	124	4	-86.74
Na^+	0.04	50	470	56.6
Cl^-	0.3	55	580	-59.51
V_m	-51.35			

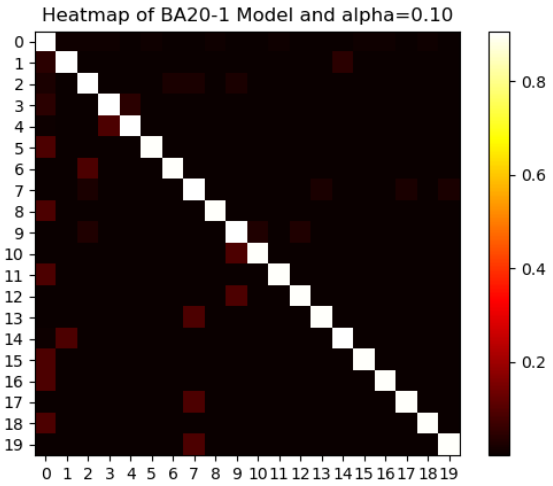

Figure 5: Heatmap. Alpha=0.1

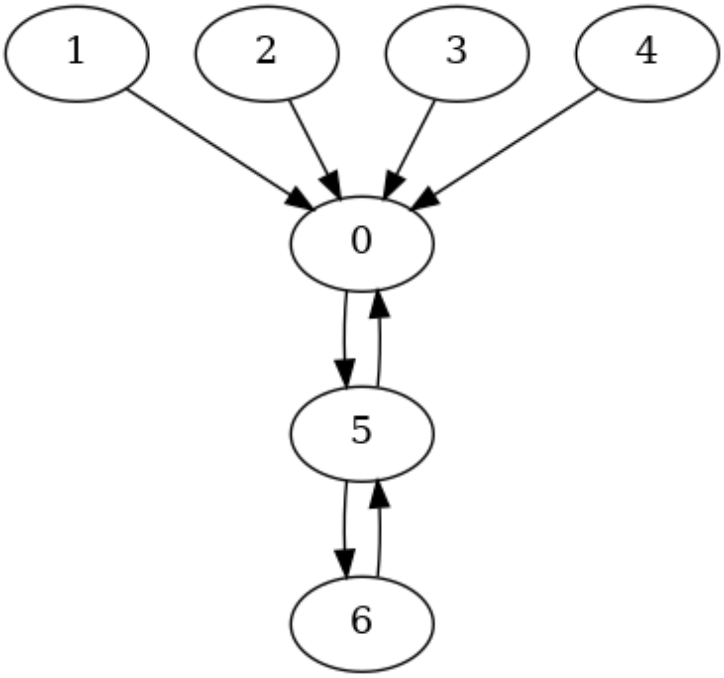

Figure 6: The apparatus used in the experiment. The forward right wing is fixed to the chain wheel which controls its elevation. The electrodes sticking out of the black precision clamp are positioned to touch the exposed meso N1 at the root of the wing

6 Reference

References

[1] Hugh Jass. "A big paper". In: *The journal of big papers* MCMXCVII (7991).

[2] Laura Lion, Gabrielle Giraffe, and Carl Capybara. *The dangers of asking the wrong question.* publishing house, 2010.

[3] Carl D Meyer. *Matrix analysis and applied linear algebra.* Vol. 71. Siam, 2000.

[4] *Perron-Frobenius Theorem.* Wikipedia. 2020. URL: `https://en.wikipedia.org/wiki/Perron%E2%80%93Frobenius_theorem` (visited on 08/30/2020).

[5] D Serre. "Matrices theory and applications second edition". In: *Graduate Texts in Mathematics* 1.216 (2010), ALL–ALL.