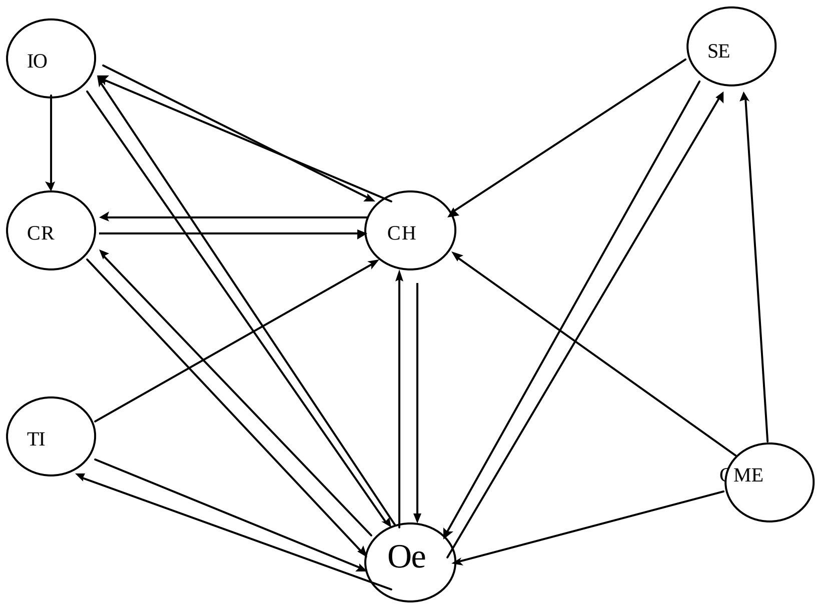
Influence Diagrams and Matrices Applied to the Determination of Character Importance in *Oedipus Rex* (a First Attempt)

Th. Kehagias
April 5, 1998

1. Introduction: This is an outline of a mathematical analysis of *Oedipus Rex* (or, for that matter, any other text equipped with characters and a plot). The goal of the analysis is to establish the (direct and indirect) influence exercised by each character in the play. The analysis is mathematical in two senses. First, it uses mathematical tools (the theory of stochastic matrices). Second, it follows the mathematical style of inference: certain assumptions are postulated; *if the assumptions are accepted*, certain conclusions can be proved to be true. Assuming that the intermediate mathematical reasoning has no faults, the conclusions are as valid (neither more nor less) as the assumptions. The evaluation of the assumptions is a matter beyond mathematical analysis and is left to each individual reader.

2. Influence Table: Before formally enunciating the principles of analysis, I present a few informal observations about *Oedipus*.

The following are the characters of the play: Chorus, Oedipus, Priest, Creon, Tiresias, Iocasta, Messenger 1, Servant, Messenger 2. I will consider that the *important* characters are: Chorus, Oedipus, Creon, Tiresias, Iocasta, Messenger 1, Servant (this restriction can be removed without substantially altering the analysis). Hence we have seven characters in the play. Assume that in some manner (more on this later) I determine the influence exercised by each character on the remaining ones. For instance, it can be argued that the servant influences Oedipus, while the converse is not true. On the other hand, the servant follows Iocasta's order to expose Oedipus, so he is influenced by her. Now Oedipus influences Iocasta and vice versa. And so on. We can represent these influences graphically, by an *influence diagram* as follows.



The arrows in the above diagram show influence from one character to another. The diagram is rather unwieldy; I have also omitted *self*-influence arrows, from each character to itself. The same information can be presented more neatly in terms of the following table.

	CH	OE	CR	TI	IO	ME	SE
CH	1	1	1	1	1	1	1
OE	1	1	1	1	1	1	1
CR	1	1	1	0	1	0	0
TI	0	1	0	1	0	0	0
IO	0	1	0	0	1	0	0
ME	0	1	0	0	0	1	0
SE	0	1	0	0	0	1	1

Table 1

Every row corresponds to a character and every position in this row to influence (or lack thereof) that the character receives from other characters. For instance, consider row SE (=Servant). The servant is influenced by his own considerations, as well as by Oedipus threats and the Messenger's exhortations. Hence in the SE row, there are ones in the OE. ME and SE columns. Since the servant has no significant interactions with the remaining characters, there are zeros in every other position in the SE row.

What good is such a diagram or table? To explain this, it must first be realized that either the diagram or the table portrays only direct (or *1st order*) influences. However, there are indirect influences. For example consider Creon. He has no first order interaction with the messenger (hence in the above table the intersection of the CR row and ME column has in it a zero. However, the messenger influences Oedipus and Oedipus certainly influences Creon; hence the messenger has a second order influence on Creon. There are also 3rd , 4th and even higher order influences.

It may be surprising, but *all* such influences (of higher orders) can be represented in terms of the above table and its properties. Before, however, embarking on an analysis of the table, I want to introduce a further refinement. So far it has been assumed that all above influences are of "equal strength". Actually this will not be true. For example, it may be argued plausibly that Oedipus receives a stronger influence from Iocasta than from the Chorus. Let us then (rather arbitrarily for the time being) assign numerical values (on a scale from zero to ten) on each character's influences. We might obtain a table such as this one.

	CH	OE	CR	TI	IO	ME	SE
CH	5	4	1	4	2	3	3
OE	3	7	3	9	6	3	5
CR	1	5	7	0	1	0	0
TI	0	3	0	9	0	0	0
IO	0	7	0	0	1	0	0
ME	0	3	0	0	0	5	0
SE	0	8	0	0	0	3	5

Table 2

I need to take one final step in the formulation of the influence table. Consider, for simplicity, the ME row. The Messenger is influenced by himself and by Oedipus. The total influence he receives is, so to speak, 8 influence units; 5 of these come from himself and 3 from Oedipus. Then we may fairly say that the messenger is *self-influenced* by 5/8=62.5%=0.625 and Influenced by Oedipus by 3/8=37.5%=0.375. OK, let's do the same trick for every row of the above table, to finally get the following *normalized influence table*.

	CH	OE	CR	TI	IO	ME	SE
CH	0.22	0.18	0.04	0.18	0.09	0.13	0.13
OE	0.08	0.19	0.08	0.25	0.16	0.08	0.14
CR	0.07	0.36	0.50	0	0.08	0	0
TI	0	0.25	0	0.75	0	0	0
IO	0	0.88	0	0	0.12	0	0
ME	0	0.38	0	0	0	0.62	0
SE	0	0.50	0	0	0	0.19	0.31

Table 3

3. The Influence Matrix and Its Properties: Here is the formal part, even though I try to express it not too formally. The *influence matrix* is basically the above table. The idea is *to use the influence matrix to determine the most influential (most significant?) character in the play.*

> *Definition*: The influence matrix P of a play with K characters is a matrix such that $P_{ij} \geq 0$ for i,j=1,2,…,K, and
>
> $P_{i1}+P_{i2}+\ldots+P_{iK}=1$ for I=1,2,…,K.
>
> The element P_{ij} expresses the proportion of influence exercised on character i by character j.•

We have already given the interpretation of the influence matrix. I will not give the full argument, but the following results is not difficult to show.

> *Lemma* 1: The element i,j of P^n (i.e. the n-th power of the influence matrix) shows the proportion of n-th order indirect influence exercised on character i by character j. •

(Here powers are expressed (as usual) in terms of multiplication, which is understood in the matrix sense.)

Hence the n-th power of the influence matrix shows n-th order interactions. Now, the definition of P guarantees that it is a *stochastic* matrix. I will give no further definition; just consider (in the context of this analysi) that "stochastic" and "influence" matrices are synonyms. Stochastic matrices have been studied extensively and their properties are very well known (to mathematicians). And there is one property of the influence matrix which is particularly important in understanding the importance of each character in a play.

Before this property is expounded, however, consider one possible weakness of using the powers of the influence matrix to determines the most influential character. It is the following: a character may be be very influential for, say, interactions of the 10th until 15th order, but completely influential for interactions of the 16th to 20th order. And then perhaps for 21st to 25th order interactions he may be again influential and then uninfluential and so on. Which particular time interval should w euse to evaluate a character's influence?

As it turns this difficulty is only apparent. In fact, under appropriate conditions, a character's influence tends to stabilize as we consider interactions of higher order. This is the subject of the next theorem.

> *Definition 2*: A stochastic matrix P is called *strongly communicating* if there is a number n such that $P^n>0$, i.e. all elements of the P^n matrix are strictly greater than zero.

> *Theorem 1*: If a stochastic matrix P is strongly communicating, then there is a matrix Q such that $\lim_{n\to\infty} P^n=Q$. Furthermore, all rows of Q are equal.•

This shows that in plays where the influence matrix is strongly communicating, by considering influences of high enough order (as $n\to\infty$) we will obtain a stable picture of each character's influence. This can be obtained by looking at *any* row of the limit matrix Q: such a row shows the influence (*in the long run*) exercised on the corresponding character by each character in the play. Two things are remarkable.

1. First, that the influence exercised on a character by another character stabilizes in the long run ($\lim_{n\to\infty} P^n = Q$.).
2. Second, that in the long run the influence exercised by a character on all other characters is the same ("all rows of Q are equal").

How can we establish that a particular matrix is strongly communicating? We can of course keep computing powers P^n for various values of n until we find some n which yields $P^n>0$. But this may take a very long time. On the other hand, we can use the following theorem.

Theorem 2: If a stochastic matrix P satisfies the following:

1. for some i we have $P_{ii}>0$;
2. for *any* pair i, j there are numbers k, l, m, …, s, t such that $P_{kl}>0$, $P_{lm}>0$,…, $P_{st}>0$;

then P is strongly communicating.•

It is worthwhile pointing out the interpretation of the above conditions in the context of an influence matrix. The first condition requires that at least one character influences his or herself. The second condition requires that every character can influence (perhaps indirectly) every other character. These are both conditions reasonable to expect in a play.

4. The Influence Matrix of *Oedipus Rex*: Let us now apply the previous results to *Oedipus*. Recall that our goal is to determine the most influential character in the play. To do this, we need the following.

1. Determine an influence matrix.
2. Establish that the influence matrix is strongly communicating.
3. Find the limiting matrix Q.
4. Taking any row of Q, find the element which has highest value; this corresponds to the most influential character.

Let us go through the list.

4.1. Determining the Influence Matrix. Let us use the influence matrix that corresponds to Table 3 (a more exact influence matrix could be determined by more detailed reference to the text).

4.2. The Influence Matrix is Strongly Communicating. In this case we can use Theorem 2. Oedipus influences everybody (including his self) and it can be checked that every character can influence (perhaps indirectly) Oedipus. So both of the conditions of Theorem 2 are satisfied.

4.3. Finding Q. This requires numerical computation. I did this using a computer and I got that the limiting influence matrix Q is given by:

	CH	OE	CR	TI	IO	ME	SE
CH	0.04	0.31	0.06	0.34	0.06	0.12	0.07
OE	0.04	0.31	0.06	0.34	0.06	0.12	0.07
CR	0.04	0.31	0.06	0.34	0.06	0.12	0.07
TI	0.04	0.31	0.06	0.34	0.06	0.12	0.07
IO	0.04	0.31	0.06	0.34	0.06	0.12	0.07
ME	0.04	0.31	0.06	0.34	0.06	0.12	0.07
SE	0.04	0.31	0.06	0.34	0.06	0.12	0.07

Table 4

4. 4 Finding the most influential character. It can be seen from Table 4 that maximum influence *on every character* is exerted by Tiresias: he exerts 34% of all the influence exerted in every character, as compared to

Oedipus who exerts 31% of all the influence exerted in every character. Ergo: Tiresias is the most influential (and most significant?) character in the play.

5. Discussion. Of course several objections can be raised to the above analysis. I will try to answer these in our next meeting.

It is important to note that the conclusion regarding Tiresias' influence is valid only to the extent that one accepts the quantitative description of influence in terms of an influence matrix, as well as the particular influence values assigned to the characters. Accepting or not the first is really an axiomatic choice. Accepting the particular values can be supported by textual analysis. There is scope here for statistical analysis of interaction between characters, e.g. counting the instances in which one character influences another character's action or opinion. Obviously, this can be done more reliably in longer text.

However, it must be noted that there is a significant robustness of the conclusion with respect to the particular values and this can be phrased in rather exact terms. In other words, arguments can be made of the form: " if the relative influences of Oedipus and Tiresias do not fluctuate by more than a given factor, then Tiresias will remain the most significant character…" and so on.

One important potential application of influence matrices in textual interpretation is in determining *disconnected* groups of characters, i.e. groups which in the long run do not influence each other. Such situations occur in the case of influence matrices which are not strongly communicated.

Another aspect not touched here (but which can be treated within the context of influence matrices) is the time variation of influence, i.e. the fact that a character's position can be strengthened or weakened as the plot evolves.

And more …

Influence Matrices in *Oedipus Rex*

Th. Kehagias
April 21, 1998

1. What can math do for us in interpreting a text? Here is a modest task: determine who is the most *influential* character in Oedipus. Perhaps this character is also most significant? I got the idea starting from Oedipus complex.

2. The model: influence diagram, 0/1 influence matrix.

	CH	OE	CR	TI	IO	ME	SE
CH	1	1	1	1	1	1	1
OE	1	1	1	1	1	1	1
CR	1	1	1	0	1	0	0
TI	0	1	0	1	0	0	0
IO	0	1	0	0	1	0	0
ME	0	1	0	0	0	1	0
SE	0	1	0	0	0	1	1

Table 1

3. Refined model: influence matrix with quantized influence.

	CH	OE	CR	TI	IO	ME	SE
CH	5	4	1	4	2	3	3

OE	3	7	3	9	6	3	5
CR	1	5	7	0	1	0	0
TI	0	3	0	9	0	0	0
IO	0	7	0	0	1	0	0
ME	0	3	0	0	0	5	0
SE	0	8	0	0	0	3	5

Table 2

4. Refined model: percentized influence matrix; now I had a math theory to work with.

	CH	OE	CR	TI	IO	ME	SE
CH	0.22	0.18	0.04	0.18	0.09	0.13	0.13
OE	0.08	0.19	0.08	0.25	0.16	0.08	0.14
CR	0.07	0.36	0.50	0	0.08	0	0
TI	0	0.25	0	0.75	0	0	0
IO	0	0.88	0	0	0.12	0	0
ME	0	0.38	0	0	0	0.62	0
SE	0	0.50	0	0	0	0.19	0.31

Table 3

5. Conclusion: under certain conditions, counting not only direct but also indirect influences, there is a limiting set of influence values: every character exercises the same influence on all characters

> *Theorem*: If a stochastic matrix P is strongly communicating, then there is a matrix Q such that $\lim_{n\to\infty} P^n = Q$. Furthermore, all rows of Q are equal.•

6. Examples …

7. OK, what does this show us? We learned that there will be a most influential character, and we can find him/her. But who it is will depend on the influence assignments. Basically we do not get anything new. Two interesting results may occur:

   7.1 Obtain our conclusions more unequivocally (no chance for fuzzy arguments)
   7.2 Obtain surprising results (counter intuitive?). In this case, it may be said that we obtain something new, or more clarified.

8. Can we find a dependable procedure for assigning influence weights?

   8.1 By asking an expert
   8.2 By asking many people
   8.3 By asking many experts
   8.4 By asking two conflicting groups of experts (schools?)

9. Mathematically interesting:

   9.1 if we ask many people, how to obtain group rankings?

9.2 Are absolute values of influence weights or should we use only an ordering?

9.3 Establish that the results are robust in some sense (so exact values do not matter very much).

10. Finally this is another case of modelling. The hard part is to obtain the correct model and good estimates of the parameters. Model building is a combination of mathematical skills and expertise. Parameter value assignment is (in my opinion) purely expertise.
11. What does all this do for proof?