Mathematical Decision Making

An Overview of the Analytic Hierarchy Process

Steven Klutho

5/14/2013

Contents

1 Acknowlegdments

I would like to thank professors David Guichard and Barry Balof for their invaluable help and guidance throughout this project. I would also like to thank Mzuri Handlin for his help reviewing and editing my work. Last but certainly not least, I would like to thank Tyler Hurlburt for being my test subject during the applied problem, and allowing me to decide his future using mathematics.

2 Introduction

Every day we are faced with decisions–what to wear, what to eat for lunch, what to spend our free time doing. These decisions are often made quickly, and in the moment. However, occasionally a decision comes around that cannot be quickly resolved. Perhaps we cannot easily see which choice is best, or maybe the repercussions of choosing wrong are much higher. It would be advantageous to have a systematic, mathematical process to aid us in determining what choice is best. This paper examines and elucidates one of these methods, the **Analytic Hierarchy Process**, or **AHP**.

There is an entire field of mathematics dedicated to decision making processes and analytics, a very important subject in the business world, where decisions involving millions of dollars are made every day. For this reason, we have to be sure the decision making processes used make sense, are efficient, and give us reasonable answers. They have to be grounded in reality, and have a method that is backed up by mathematical theory. We will show throughout this paper that the AHP is a useful, appropriate method for dealing with decision problems.

In order to understand what the AHP is capable of, we will first define just what a decision problem is. Once we have this background, we will present the tools from linear algebra that we will need going forward. This paper assumes a knowledge of basic terms from linear algebra, as well as matrix multiplication. After developing our tools, we begin to turn our attention to the specific workings of the AHP. We look at the types of problems it is designed to solve, the principles guiding the process, and the mathematics at work in determining the best decision. We will then discuss the idea of inconsistency in the AHP, that is, when we can trust the data that the process gives us. Finally, we will work through an involved, real world example of the AHP at work from start to finish.

3 Linear Algebra Review

This section is essentially a toolbox for theorems and definitions we will need later when developing the AHP. It may be skipped and referred back to as needed.

3.1 Eigenvectors and Eigenvalues

Both eigenvectors and eigenvalues are very important in how the AHP works. We define them here, and list some of their pertinent properties we will use going forward.

Definition 1. *An* ***eigenvector*** *of a square matrix A is a vector $\boldsymbol{v}$ such that*

$$A \times \boldsymbol{v} = \lambda \boldsymbol{v} \tag{1}$$

Definition 2. *An* ***eigenvalue*** *is the scalar λ associated with an eigenvector $\boldsymbol{v}$.*

Note that if we multiply both sides of equation 1 by a scalar, we get the same eigenvalue, but a different eigenvector, which is simply a scalar multiple of our original v. All scalar multiples of v will have the same eigenvalue (and so are all in the same eigenspace), and so it may be a little ambiguous to determine which eigenvector we are dealing with when we discuss a given eigenvalue. In this paper, it can be assumed unless explicitly stated, that we are looking at the eigenvector with entries that sum to 1.

3.2 Building Block Theorems for the AHP

With an understanding of eigenvectors and eigenvalues, we can now present a few definitions and theorems we will use later in the paper. We present the theorems here without proof, since presenting the proof would do little to help the understanding of the AHP, which is the goal of this paper.

Definition 3. *The* ***rank*** *of a matrix is the number of linearly independent columns of that matrix.*

Theorem 1. *[6] A matrix of rank one has exactly one nonzero eigenvalue.*

Definition 4. *The* ***trace*** *of some matrix A, denoted $Tr(A)$, is the sum of the entries in the diagonal.*

Theorem 2. *[6] The trace of a matrix is equal to the sum of its eigenvalues.*

4 Decision Problems

As alluded to in the introduction, decision problems present themselves in a wide variety of forms in our everyday lives. They can be as basic as choosing what jacket to buy, and as involved as determining which person to hire for an open position [6]. However, across these different problems, we can often present each as a *hierarchy*, with the same basic structure.

4.1 The Hierarchy

Each decision problem can be broken up into three components, defined here.

Definition 5. *The **goal** of the problem is the overarching objective that drives the decision problem.*

The goal should be specific to the problem at hand, and should be something that can be examined properly by the decision makers. For instance, in the example of hiring a new employee, different departments would have very different goals, which could steer their decision problem to different outcomes. Moreover, the goal should be singular. That is, the decision-makers should not attempt to satisfy multiple goals within one problem (in practice, these multiple goals would be broken into separate **criteria**, as defined below). An example of a goal with regards to buying a jacket could be "determining the jacket that best makes me look like a Hell's Angel", which is specific, singular, and something that can be determined by the decision maker.

Definition 6. *The **alternatives** are the different options that are being weighed in the decision.*

In the example of hiring a new employee, the alternatives would be the group of individuals who submitted an application.

Definition 7. *The **criteria** of a decision problem are the factors that are used to evaluate the alternatives with regard to the goal. Each alternative will be judged based on these criteria, to see how well they meet the goal of the problem.*

We can go further to create sub-criteria, when more differentiation is required. For instance, if we were to look at a goal of buying a new car for a family, we may want to consider safety as a criterion. There are many things that determine the overall safety of a car, so we may create sub-criteria such as car size, safety ratings, and number of airbags.

With these three components, we can create a hierarchy for the problem, where each level represents a different cut at the problem. As we go up the hierarchy from the alternatives, we get less specific and more general, until we arrive at the top with the overall goal of the problem. See Figure 1 for the layout of such a hierarchy. Note that not every criterion needs sub-criteria, nor do those with sub-criteria need the same number of sub-criteria. The benefits for structuring a decision problem as a hierarchy are that the complex problem is laid out in a much clearer fashion. Elements in the hierarchy can be easily removed, supplemented, and changed in order to clarify the problem and to better achieve the goal.

4.2 Weighting the Problem

Another component in any decision problem is a mapping of notions, rankings, and objects to numerical values [6]. Basic examples of such mappings are methods of measurement we are familiar with, such as the inch, dollar, and kilogram. These are all examples of **standard scales**, where we have a standard unit used to determine the "weight" of each measurement in the various scales. We run into problems, however, when we analyze things for which there is no standard scale. How do we quantify comfort? We have no set "cushiness scale" to compare one sofa to another. However, given two sofas to compare, in most cases we will be able to determine one to be more comfortable than the other. Moreover, we can often give a general value to how much more comfortable one is than the other. In making pairwise comparisons between different options, we can create a **relative ratio scale**, which is the scale we will use when dealing with the AHP.

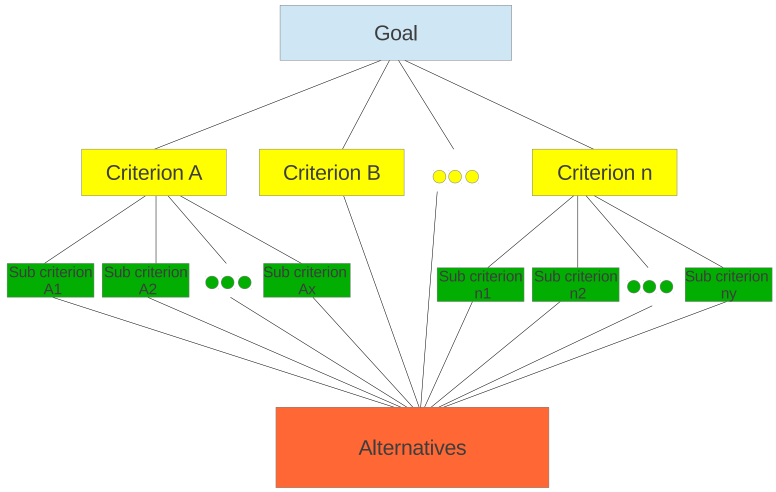


Figure 1: A generic hierarchy

With this idea of a relative ratio scale in mind, let us reexamine the idea of a standard scale. Consider the Celsius scale, which is certainly a standard scale. Suppose now we are looking at a decision problem with the goal of determining the optimal temperature to set the refrigerator. Certainly a temperature of 100 degrees would be foolish, as would a temperature of -100 degrees. In this case, our standard scale does us no good. An increase in the "weight" of degrees Celsius does not necessarily correspond to a better temperature for reaching our goal. In this case, our standard scale turned into a relative ratio scale, given the goal of the problem. Another very important example to note is the standard scale of money [6]. Often, economists assume the numerical value of a dollar to be the same regardless of the circumstances. However, when looking at buying a new yacht, \$100 is just about as useless as \$10, despite being ten times as much money. However, when buying groceries, \$100 suddenly becomes a much better value, perhaps even surpassing its arithmetic value when compared to \$10. In this case, we have greatly increased our spending power, whereas when looking at the yacht, we have done almost nothing to increase it. We may increase our resources tenfold, but we have had no real affect on our spending power.

We can see that the way we weight a decision problem is very important. Furthermore, in most circumstances, standard scales will do us no good. Hence, we will have to resort to pairwise comparisons, and a relative ratio scale that we determine from these comparisons.

5 How the AHP Works

This section will deal with the ideas at work behind the AHP, the way rankings are calculated, and how the process is carried out.

5.1 Creating the Matrices

Recall the ideas presented in Section 4.2 about standard and ratio scales. If it were the case that we were dealing with a standard scale, we would have a set of n objects with weights $w_1, w_2, \ldots, w_n$. We could then create a matrix of comparisons, A, giving the ratio of one weight over another, as shown in equation 2.

$$A = \begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \ldots & w_n/w_n \end{bmatrix} \tag{2}$$

This matrix is an example of a *consistent matrix*, which we will define formally here.

Definition 8. *A* ***consistent matrix*** *is one in which for each entry a_{ij} (the entry in the i^{th} row and j^{th} column), $a_{ij} = a_{kj}/a_{ki}$.* [6]

What this implies in our matrices is that for a consistent matrix there is some underlying standard scale. That is, each element has a set weight, which does not change when compared to another element. Hence, we are able to do the calculation described in Definition 8, and we will have cancellation.

Note that the ratio in entry a_{ij} is the ratio of w_i to w_j. This will be the standard we adopt for the rest of the paper, where an entry in a ratio matrix is to be read as the ratio of the row element to the column element.

Now that we have our ratio matrix, notice that we can create the following matrix equation:

$$\begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \ldots & w_n/w_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \tag{3}$$

We know that the matrix equation in 3 is valid because when we go through the matrix multiplication, we see that we have an $n \times 1$ matrix with entries:

$$w_1(w_i/w_1) + w_2(w_i/w_2) + \cdots + w_n(w_i/w_n)$$

for row i of the product matrix. As can be seen, the denominators cancel, and we are left with nw_i for row i. Hence, we can factor out the scalar n and express the product matrix as we have in equation 3. We can see that n is then an eigenvalue for the $n \times n$ ratio matrix (which we will call matrix A), and that the $n \times 1$ weight matrix (which we will call matrix W) is an eigenvector. We know because of Theorem 2 that the sum of the eigenvalues of matrix A are simply its trace. Since

$$Tr(A) = (w_1/w_1) + (w_2/w_2) + \cdots + (w_n/w_n) = 1 + 1 + \cdots + 1$$

we know that $Tr(A) = n$. Hence, we know that the sum of the eigenvalues of matrix A is equal to n. Since we have already shown n to be an eigenvalue of A, and since matrix A is of rank one, we know by Theorem 1 that n is the only nonzero eigenvalue. We now make the following definition.

Definition 9. *The* ***principal eigenvalue****, denoted λ_{max}, of a matrix is the largest eigenvalue of that matrix.*

As can be seen, for a standard scale ratio matrix $\lambda_{max} = n$. However, we are not very interested in these standard scales, since they often do not tell us much about the real world problems we will be dealing with. Instead, we will use the ideas behind what we just did to look at a similar problem with a relative ratio scale. We are much more interested in the ratio matrices created when we look at comparisons between things in a relative ratio scale setting, since these are the kinds of scales typical of decision problems. By making pairwise comparisons between alternatives, we can easily construct a ratio matrix similar to the one in equation 2. The result is shown here.

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 1/a_{12} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & ... & 1 \end{bmatrix} \tag{4}$$

Note that we have reciprocal values across the diagonal, since we are simply inverting the ratio. Furthermore, we have all ones down the diagonal since comparing an alternative to itself would result in a ratio of 1 : 1. This matrix is an example of a *reciprocal matrix*, which we define here.

Definition 10. *A* ***reciprocal matrix*** *is one in which for each entry a_{ij}, $a_{ji} = 1/a_{ij}$.*

With regards to the ratio matrices we construct, the fact that they are reciprocal stems from the fact that the ratio does not change depending on which element you compare to another. We assume that comparing option A to option B is the reciprocal value of comparing option B to option A. Note that every consistent matrix is a reciprocal matrix, but not all reciprocal matrices are consistent. However, what if the ratio matrix given in equation 4 was consistent? We would then have an underlying standard scale, and we would be able to create a ranking of the elements based on this scale. Thus, let us construct a similar matrix equation as that given in equation 3.

$$\begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 1/a_{12} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & ... & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \lambda_{max} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \tag{5}$$

The eigenvector corresponding to λ_{max} in this equation is essentially our underlying standard scale, and thus gives us the ranking of each element in the ratio matrix. Hence, determining the rankings for a set of elements essentially boils down to solving the eigenvector problem

$$AW = \lambda_{max} W$$

where W is the weight matrix of the alternatives in question. This is, in essence, the principle that the AHP works on–that given some group of elements, there is an underlying standard scale. Each element has a numerical value in this scale, and can thus be compared numerically with other elements in the group. With the weight matrix in equation 5, we could create a new ratio matrix based only on the weights that come from solving the eigenvector problem. This matrix would then be consistent, and if there was indeed some kind of underlying scale, it should have entries very close to those in the original ratio matrix. If this new matrix was too inconsistent to give a reliable weight matrix, then it did not have much of an underlying scale to begin with. See Section 7 for a discussion on determining if the original ratio matrix is close enough to this new ratio matrix, and what to do if it is not.

5.2 Determining the Weights

Now that we have an idea of how to determine the weight given a ratio matrix, we have to translate that skill into a real world problem. Specifically, we want to determine the overall weight for every alternative in a given decision problem. In order to show how we do this, we will work through the very simple example of buying new computers for a college lab.

First, we determine the goal of our decision problem. In this case, it is simply to choose the best computer for a mathematics computer lab at Whitman College. Then, we determine the criteria we will use to see how each computer meets the goal. In this example, we will use the following criteria:

- CPU
- Cost
- Looks
- What professor Schueller wants

We could certainly add more to this list, but for this example, we decide these are the factors we most want to look at in order to achieve the goal. We could even add more sub-criteria, creating a larger hierarchy (for instance, under the "looks" criterion, we could have the sub-criteria "shininess" and "color"). Finally, we determine our alternatives. In this simple example, we will simply consider three computers, X, Y, and Z. Given these elements of our hierarchy, we can create the visual representation given in Figure 2.

We must now determine the weight for each criterion. These values will tell us just how important each criterion is to us in our overall goal. To determine these values, we make pairwise comparisons between the criteria and create a ratio matrix as in equation 4, using the numerical values given in Table 15. In making these comparisons, it is important to only consider the two criteria being compared. We can then begin to size up how important each criterion is to us.

Obviously what professor Schueller wants is vital, as is how the computer looks, because everyone knows that is what is most important. Let's just say we're loaded, so cost is of less importance, and CPU takes a position somewhere in the middle of the road. With these general rankings in mind, we then compare each criterion pairwise with each other criterion, disregarding the others for the moment. It is important to note that if we assign a value comparing some criterion to another, then the comparison going the other way will have a reciprocal value. For instance, if we decide that CPU has moderate importance over cost, we may assign a value of 3 to this comparison. Thus, the value of cost compared to CPU would be $\frac{1}{3}$.

Hence, we can construct a 4×4 matrix of these values, where the criterion on the left side of the matrix is the first criterion in the comparison. Thus, for example, we could have a matrix that looks

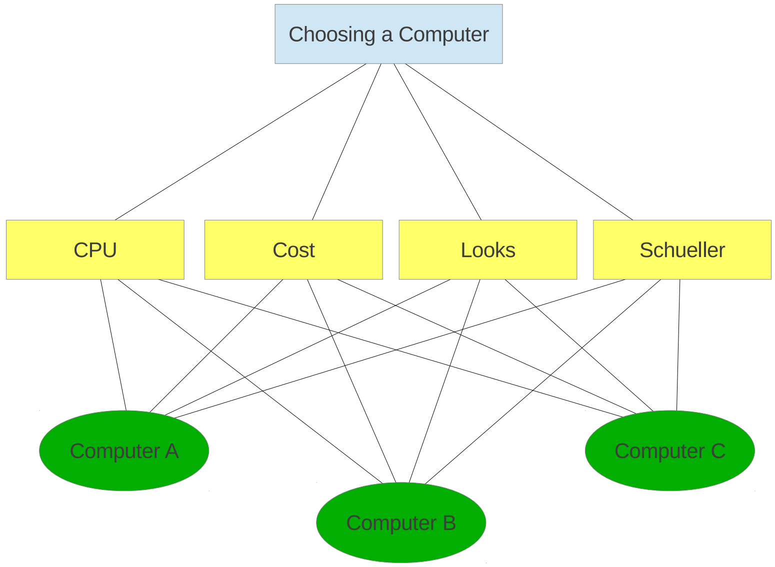


Figure 2: The Hierarchy

something like Table 1. From this matrix, we can then calculate λ_{max} and its associated eigenvector. Normalizing this eigenvector gives us the matrix W_c, which shows the relative "weights" of each criterion, determining how much sway they have in determining what the eventual choice will be. The matrix W_c is given in Table 2. Note that normalizing does not change the ratio matrix or the eigenvalue, since we are interested only in the ratio of one weight to another.

Table 1: The ratio matrix of criteria comparisons

	CPU	Cost	Looks	Schueller
CPU	1	3	$\frac{1}{4}$	$\frac{1}{6}$
Cost	$\frac{1}{3}$	1	$\frac{1}{6}$	$\frac{1}{8}$
Looks	4	6	1	$\frac{1}{2}$
Schueller	6	8	2	1

We then must calculate the weight of each alternative in regard to each criterion. That is, we must determine how each computer stacks up in each category. Thus, we must create a separate ratio matrix for every criterion. When creating these matrices, we are only interested in how each computer compares to the other within the criterion in question. That is, when we create the

Table 2: Criteria weight matrix

Criterion	**Weight**
CPU	0.103
Cost	0.050
Schueller	0.315
Looks	0.532

"looks" ratio matrix for the alternatives, we will be comparing them based solely on their looks and nothing else. Doing similar calculations to those we did when determining the criteria weight matrix, we can put forth the following weight matrices, conveniently put together into one table, Table 3.

Table 3: Weights of Computers X, Y, and Z

	CPU	Cost	Looks	Schueller
X	0.124	0.579	0.285	0.310
Y	0.550	0.169	0.466	0.406
Z	0.326	0.252	0.249	0.285

Again, these weights were created through pairwise comparisons using the Fundamental Scale (see Table 15), and so we are dealing with a relative ratio scale. We are trying to assess how important each criterion is with regards to our problem, and then how well each alternative satisfies each criterion. With the weights we just found, we can then determine how well each alternative stacks up given the original goal of the problem. To do this, we multiply each alternative's weight by the corresponding criterion weight, and sum up the results to get the overall weight. The calculation is presented here.

$$\begin{aligned}
X_{score} &= (0.103 \times 0.124) + (0.050 \times 0.579) + (0.315 \times 0.285) + (0.532 \times 0.310) = 0.296 \\
Y_{score} &= (0.103 \times 0.550) + (0.050 \times 0.169) + (0.315 \times 0.466) + (0.532 \times 0.406) = 0.429 \\
Z_{score} &= (0.103 \times 0.326) + (0.050 \times 0.252) + (0.315 \times 0.249) + (0.532 \times 0.285) = 0.276
\end{aligned}$$

This calculation gives us the overall weight of each alternative. These weights then tell us which alternative will best achieve our goal. As can be seen, computer Y is the clear best choice, and is thus the computer that best achieves our goal.

As can be seen, the AHP is a powerful tool for solving decision problems that can be broken down into a hierarchy. This example was very rudimentary, but the process can be expanded to include many more criteria, levels of sub-criteria, and a larger number of alternatives. Moreover, we have shown that the AHP is an excellent tool for ranking very incorporeal things, such as the "looks" of a computer, and also for incorporating more standard rankings, such as "cost". We have demonstrated that there is a clear process, which results in a distinct ranking of the alternatives compared. In the following sections, we will look deeper at the theory behind the AHP, and the idea of consistency in our resulting weight matrices.

6 AHP Theory

In the previous section, we saw an example of how we can use linear algebra to try to get results that make sense. But what do these results mean? Certainly we want them to represent some kind of ranking of the alternatives, but is there any mathematical reasoning behind the calculations we did in the previous section?

To answer these questions, we will examine some of the theory at work behind the AHP. Specifically, we will look at the connection between the AHP and graph theory and the notion of dominance. We will first try to get a handle on the idea of dominance.

Intuitively, dominance is a measure of how much one alternative is "better" than another. For instance, let us consider some alternative a amongst a group of three alternatives. We can look at the ratio matrix created by comparing each alternative pairwise. The dominance of a is then the sum of all the entries in row a, normalized. For example, in Table 4, we see that the dominance of a is $1 + \frac{1}{5} + 3 = 4.2$, normalizing gives $\frac{4.2}{14.033} = 0.299$. The dominance of b and c are then respectively 0.463 and 0.238.

Table 4: Ratio matrix for alternatives a, b, and c

	a	b	c
a	1	$\frac{1}{5}$	3
b	5	1	$\frac{1}{2}$
c	$\frac{1}{3}$	2	1

This gives us a value that tells us how much a dominates the other values in a single step, which we will call the **1-dominance**. Hence, we present the following definition.

Definition 11. *The **1-dominance** of alternative i is given by the normalized sum of the entries $a_{i,j}$ where $j = 1, 2, \ldots, n$ in the ratio matrix.*

We will address the idea of other values of dominance shortly (i.e., k-dominance for some integer k), but for now this will be our working idea of dominance. This value gives us an idea of how much each alternative dominates the others. From this matrix, it is easy to see that b dominates the other alternatives most. However, let us consider the following case, where dominance is not quite as clear cut.

Consider five teams in the Northwest DIII basketball conference, namely, Whitman, Lewis and Clark, George Fox, Linfield, and Whitworth. We want to see which team is the best in the conference, and so we will consider games from the 2012-2013 season, looking at the most recent match-ups of each of these teams. We can represent the results of the season in a matrix, where a 1 in entry a_{ij} corresponds to team i defeating team j, and a 0 corresponds to team i losing to team j. We will choose to put ones along the diagonal, since we want all teams to be winners so they aren't sad (we can choose ones or zeros along the diagonal. Mathematically, it will make no difference in our final outcome). Thus, we have a matrix as in 5.

Table 5: Conference basketball wins

	Whtmn	L&C	GF	Lin	Whtw
Whtmn	1	1	1	1	1
L&C	0	1	1	1	0
GF	0	0	1	1	1
Lin	0	0	0	1	0
Whtw	0	1	0	1	1

Table 6: Weighted conference basketball wins

	Whtmn	L&C	GF	Lin	Whtw
Whtmn	1	1.042	1.208	1.431	1.033
L&C	0.960	1	1.014	1.356	0.932
GF	0.828	0.986	1	1.364	1.099
Lin	0.700	0.737	0.733	1	0.616
Whtw	0.968	1.073	0.910	1.623	1

As we can see, Lewis and Clark, George Fox, and Whitworth all have a dominance value of 0.2. How do we decide which team did better overall in the conference? The problem becomes even more difficult when we notice that Lewis and Clark beat George Fox, who beat Whitworth, who beat Lewis and Clark. Thus, in order to get a clearer idea of the problem, will consider the “weighted” wins of each team, given in Table 6. This matrix was created by looking at the ratio of each team’s final score in the given match up. Thus, we can get an idea of how much each team dominated the others in the match ups. For instance, in the game between Whitman and Whitworth, we have a value of 1.033 (0.968 if we look at it from the standpoint of Whitworth vs. Whitman). This corresponds to a much closer game than the Whitman vs. Linfield game, which has a point ratio of 1.431. In the first instance, Whitman scored 1.033 points for each point Whitworth scored, and in the second instance, Whitman scored 1.431 points for each Linfield point.

By looking at the values in Table 6 and calculating just as we did for Table 4, we find the dominance given in Table 7.

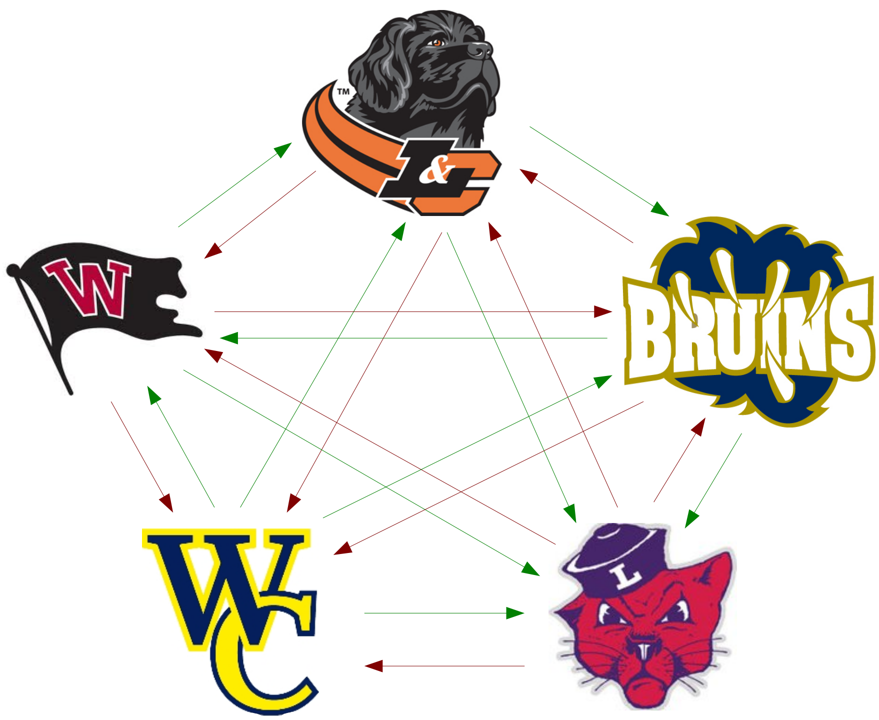
Let us look at the graph of this situation to try to get a better idea of the problem at hand,

Table 7: 1-dominance rankings

Whtmn	0.223
L&C	0.205
GF	0.206
Lin	0.148
Whtw	0.218

and to see where the idea of dominance from graph theory comes into play. The graph of this problem is given in Figure 3. Here, each vertex represents a team, and the edges represent a game being played between teams. A green arrow pointing to a team indicates a win over that team. For example, the green arrow pointing from Whitman to Linfield indicates a win for Whitman over Linfield. This arrow can then be "weighted" by assigning it the corresponding value of 1.431 from Table 6. Similarly, a red arrow indicates a loss, and is weighted as well. Now we can see that to calculate the dominance for some vertex we simply add up the weights of all the edges leaving that vertex and normalize.

Figure 3: The graph of the conference games



With the idea of calculating the dominance from the graph in mind, let us look at the edge from the Lewis and Clark vertex to the George Fox vertex. It is green, and corresponds to a win, so in our 1-dominance frame of mind, we would say that Lewis and Clark dominated George Fox. However, let us consider a walk from George Fox to Lewis and Clark, following only green arrows. We can see that such a walk is possible by going from George Fox to Whitworth to Lewis and Clark. Thus, in a way, George Fox dominates Lewis and Clark. This idea is known as *two-step* dominance. We can calculate this two-step dominance value by multiplying the weight of the GF $\rightarrow$ Whtw edge by the weight of the Whtw $\rightarrow$ GF edge.

To get a clearer picture of who really is the best team in the conference, we would probably want to consider these two-step dominance values as well. To calculate the total two-step dominance of each alternative (i.e., the normalized sum of all these weighted walks of length two beginning at a given vertex) we can square the original matrix given in Table 6 and sum across the rows [5]. Normalizing these values gives us the two-step dominance values in Table 8

Table 8: 2-dominance rankings

Whtmn	0.224
L&C	0.206
GF	0.206
Lin	0.148
Whtw	0.216

So, if we want to view a more complete picture of which alternative dominates which, we may want to consider two-step dominance in addition to one-step dominance. One way to do this would be to add these two values together, and then average them. But why stop at two-step dominance? We may want to continue further, including all the way up through some k-dominance. We could then extend even further, including every dominance step up to ∞-dominance. We then present the following theorem.

Theorem 3. *The dominance of each alternative up through ∞-dominance is given by the solution of the eigenvalue problem $AW = \lambda_{max}W$, where the entries in W are the weights of ∞-dominance corresponding to each alternative.*

Notice that this is the exact eigenvalue problem we are tackling with the AHP. From this theorem, we can see that the AHP is giving us a ranking based on the ∞-dominance of the alternatives. To prove this result, we need the following definition.

Definition 12. *[6] The dominance matrix of a set of alternatives up through some k-dominance, $k \leq m$ is given by*

$$\frac{1}{m}\sum_{k=1}^{m}\frac{A^k e}{e^T A^k e} \tag{6}$$

where $e = [1\ 1\ \ldots\ 1]^T$, and A is the ratio matrix of the alternatives.

Notice that $A^k e$ is a matrix of the row sums of A^k, and that $e^T A^k e$ is the sum of all the entries of A^k. Hence, all we are doing is finding the matrix of dominance values for each $i \leq k$, normalizing each one, and then adding them all together. We then divide this matrix by the number of matrices we added together, essentially finding an average value of dominance for each alternative.

What Theorem 3 tells us is that we can calculate the weight matrix W by repeatedly raising matrix A to increasing powers, normalizing, and then adding up all of our results. We can thus calculate our weight matrix by following this process to some k value, where the difference in the entries of matrix W resulting from $k+1$ is less than some predetermined value when compared to the W generated by k.

We will also need the following theorems.

Theorem 4. *[4] Let A be a square matrix with all entries $a_{ij} > 0$. Then*

$$\lim_{m\to\infty} (\rho(A)^{-1}A)^m = L \tag{7}$$

where $\rho(A)$ is defined as the maximum of the absolute values of the eigenvalues. Furthermore, we have the following stipulations: $L = xy^T$, $Ax = \rho(A)x$, $A^Ty = \rho(A)y$, all entries in vectors x and y are greater than 0, and $x^Ty = 1$. The vectors x and y are any two vectors that satisfy the given conditions.

Theorem 5. *[3] Let s_n be a convergent sequence, that is, let*

$$\lim_{n\to\infty} s_n = L$$

Then

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} s_k = L \tag{8}$$

Thus, we proceed with the proof.

Proof. We can see that the $\rho(A)$ in our case will simply be λ_{max}, and thus the matrix x in Theorem 4 must be an eigenvector associated with λ_{max}. Hence, we know that x is some multiple of the normalized W matrix. With this in mind, let us examine the expression within the summation in equation 6 as $m \to \infty$. We can multiply the limit by one to get the following result:

$$\lim_{m\to\infty} \frac{\frac{1}{\lambda_{max}^m}}{\frac{1}{\lambda_{max}^m}} \frac{A^m e}{e^T A^m e} = \lim_{m\to\infty} \frac{\frac{A^k}{\lambda_{max}} e}{e^T \frac{A^k}{\lambda_{max}} e}$$

We know by Theorem 4 that the $\frac{A^k}{\lambda_{max}}$ terms are equal to xy^T, as defined in that theorem. Hence, we have:

$$\lim_{m\to\infty} \frac{xy^Te}{e^Txy^Te}$$

Note that both x and y are vectors, and so y^Te is a scalar, and we can cancel this scalar from both numerator and denominator. Hence, we are left with the following result:

$$\frac{x}{e^Tx}$$

Examining the e^Tx term, we find that this term is simply a scalar equal to the sum of the entries in the x vector. Hence, dividing by this scalar normalizes the vector, giving us W. Then, by Theorem 5, we can see that when we look at the average of the summation of these terms, we get the desired result.

□

Thus, we now have a theoretical backing to our process. That is, the rankings given by the AHP actually show the ∞-dominance of the alternatives. We can see that the AHP is not just a useful trick of linear algebra, but that it is actually calculating the ∞-dominance of all of the alternatives.

To give another example of the graphical idea of dominance and how it pertains to decision problems, let us again consider the decision problem given in Section 5.2. We can then create a

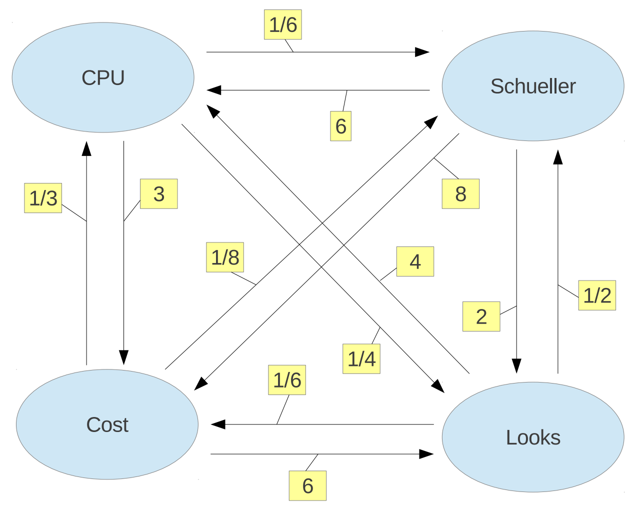


Figure 4: A graphical representation of dominance

corresponding directed graph where each vertex is a criteria, and each edge represents a comparison. Thus, an edge running from CPU to Cost takes on the weight 3, whereas the edge running the other direction has weight $\frac{1}{3}$. The graph of this example is shown in Figure 4.

We can then think of the weight of the edge as corresponding to dominance. Hence, a one-step dominance value for Cost is the sum of all the edges emanating from the Cost vertex. This value would then be

$$\frac{1}{6} + \frac{1}{3} + \frac{1}{8} = \frac{5}{8}$$

To find a two step dominance value, we would start at the Cost vertex and then move to another vertex. This corresponds to one step. Then, from that vertex, we would follow another edge, multiplying this edge's value by the previous edge. An example of one of these walks would be taking Cost to Looks, and then Looks to CPU. The value of this two-dominance would then be

$$\frac{1}{6} \times 4 = \frac{2}{3}$$

Adding up all such walks through the graph, we find a value for the two-step dominance of Cost. Similarly, we can follow the same procedure for calculating the k-step dominance.

7 Consistency in the AHP

Our whole method for determining the rankings of the alternatives is based on the idea that they have some underlying scale. As discussed earlier, this essentially boils down to the idea that when we have calculated our weight matrix, a consistent ratio matrix made using these weights isn't too far off of our original ratio matrix. Thus, in order to determine if our results are even valid, we have to come up with some way of measuring how far off we are. That is, how inconsistent our ratio matrices end up being.

7.1 The CR, CI, and RI

In this section we present the method developed by Saaty [6] for determining inconsistency. To create the tools we need to analyze the inconsistency, we will first need to prove two theorems.

Theorem 6. *[6] For a reciprocal $n \times n$ matrix with all entries greater than zero, the principal eigenvalue, λ_{max} will be greater than or equal to n. That is, $n \leq \lambda_{max}$.*

Proof. Consider some ratio matrix A, and its corresponding consistent ratio matrix A', which is created using the resulting weight eigenvector calculated from A. We know that matrix A has entries

$$a_{ij} = (1 + \delta_{ij})(w_i/w_j)$$

where $1 + \delta_{ij}$ is a perturbation from the consistent entry. We know that $\delta_{ij} > -1$ since none of our ratio matrices will ever have a negative value or 0 as an entry. (This works out well for the ratio matrices we will use, since there is no notion of a "negative preference" in the Fundamental Scale–the lowest we can go is $\frac{1}{9}$). We also know that our matrix A is a reciprocal matrix, and thus for entry a_{ji}, we have:

$$a_{ji} = \frac{1}{1 + \delta_{ij}}(w_j/w_i)$$

Thus, setting up our matrix equation $AW = \lambda_{max}W$, we have matrix A as:

$$A = \begin{bmatrix} 1 & (1+\delta_{12})(w_1/w_2) & (1+\delta_{13})(w_1/w_3) & \cdots & (1+\delta_{1n})(w_1/w_n) \\ \frac{1}{1+\delta_{12}}(w_2/w_1) & 1 & (1+\delta_{23})(w_2/w_3) & \cdots & (1+\delta_{2n})(w_2/w_n) \\ \frac{1}{1+\delta_{13}}(w_3/w_1) & \frac{1}{1+\delta_{23}}(w_3/w_2) & 1 & \cdots & (1+\delta_{3n})(w_3/w_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{1+\delta_{1n}}(w_n/w_1) & \frac{1}{1+\delta_{2n}}(w_n/w_2) & \frac{1}{1+\delta_{3n}}(w_n/w_3) & \cdots & 1(w_n/w_n) \end{bmatrix}$$

Since we know matrix A is a reciprocal matrix, we know that the terms in the diagonal must all be 1.

We then multiply this matrix by W, which has entries $w_1, w_2, ..., w_n$, as in the matrix equation. The result is equal to $\lambda_{max}W$. For each element in this matrix, we can look to see that it is equal to the sum of the terms in the corresponding row in matrix A, without the w_j in the w_i/w_j term. Hence, we can divide both sides by w_i, and we have an equation giving us λ_{max} for each row. Adding up all of these rows, we have something that is equal to $n\lambda_{max}$, where each $(1 + \delta_{ij})$ occurs twice–once normally, and once as its inverse. Hence, we can find a common denominator and find what these terms look like in the sum:

$$\begin{aligned}
(1+\delta_{ij}) + \frac{1}{1+\delta_{ij}} &= \frac{(1+\delta_{ij})^2}{(1+\delta_{ij})} + \frac{1}{1+\delta_{ij}} \\
&= \frac{2+2\delta_{ij}+\delta_{ij}^2}{1+\delta_{ij}} \\
&= \frac{2(1+\delta_{ij})+\delta_{ij}^2}{1+\delta_{ij}} \\
&= 2 + \frac{\delta_{ij}^2}{1+\delta_{ij}}
\end{aligned}$$

Thus, summing up all of these terms, we have

$$n + \sum_{1\leq i<j\leq n} 2 + \frac{\delta_{ij}^2}{1+\delta_{ij}} = n\lambda_{max}$$

The summation comes from the fact that we don't want to count terms twice, so we just sum over the elements in the super diagonal and up, which actually gets us terms from the entire matrix except for the diagonal. The sum of these diagonal terms is n. Since we are thus summing over $\frac{n^2-n}{2}$ elements, the 2 inside the sum can be pulled out to create the equality

$$\begin{aligned}
n + n^2 - n + \sum_{1\leq i<j\leq n} \frac{\delta_{ij}^2}{1+\delta_{ij}} &= n\lambda_{max} \\
n + \frac{1}{n}\sum_{1\leq i<j\leq n} \frac{\delta_{ij}^2}{1+\delta_{ij}} &= \lambda_{max} \\
\frac{1}{n}\sum_{1\leq i<j\leq n} \frac{\delta_{ij}^2}{1+\delta_{ij}} &= \lambda_{max} - n
\end{aligned}$$

And thus, because of the way we defined δ_{ij}, we know that the left side of the equation is greater than zero, and thus, we have shown that $\lambda_{max} \geq n$ □

Theorem 7. *A is consistent if and only if* $\lambda_{max} = n$

Proof. If A is consistent, then we know by definition 8 that each row of A is a multiple of the first row, and thus we know that A has rank 1. Hence, we then know that A has only one nonzero eigenvalue. We also know by the definition of consistent that the terms of A along the diagonal must be 1, and therefore $Tr(A) = n$. Since we have shown that the sum of a matrix's eigenvalues is equal to its trace, A must therefore have $\lambda_{max} = n$.

Now suppose A has $\lambda_{max} = n$. By Theorem 6, we know that each δ_{ij} must then be equal to 0. Hence, we have matrix A of the form in equation 2, which we know to be consistent. □

With these two theorems, we can now tackle the idea of measuring inconsistency. We first define the *consistency index*.

Definition 13. *The **consistency index (CI)** is the value*

$$\frac{\lambda_{max} - n}{n - 1} \tag{9}$$

Let us examine what this value actually measures. We know from our theorems that an inconsistent matrix will have a principal eigenvalue greater than n. Hence, the numerator is essentially measuring how far off the principal eigenvalue is from the consistent case. We then divide this value by $n-1$, which gives us the negative average of the other $n-1$ eigenvalues of our matrix [6]. With this idea of a consistency index, we can extend further to look at what the CI would be for some completely random reciprocal matrix. Moreover, we can find the CI for a large number of these matrices, and determine an average "random" CI. We then have the following definition.

Definition 14. *The **random index (RI)** of size n is the average CI calculated from a large number of randomly generated reciprocal matrices.*

In our case, the RI will be determined from 500,000 matrices, randomly generated by computer. For a discussion as to how these matrices were created, see [1]. Given these two definitions, we can make a third, which will give us a value of the inconsistency.

Definition 15. *The **consistency ratio (CR)** of a reciprocal $n \times n$ matrix A is the ratio of $CI(A) : RI(A)$, where $RI(A)$ is the random index for matrices of size n.*

This consistency ratio tells us essentially how inconsistent our matrix is. When running through the AHP, if we ever encounter a CR greater than 0.1, we have a ratio matrix that is too inconsistent to give reliable results [6]. In this case, there is not much of an underlying scale present, and the process does not work. Thus, if we ever encounter a CR greater than 0.1, we will go back to the ratio matrix and try to reevaluate our comparisons, until we get a CR that falls within our parameters. For the purposes of this paper, when we encounter a matrix that falls within this tolerance, we will refer to it as "consistent enough." Similarly, a matrix that falls outside out tolerances is a matrix which is "too inconsistent."

7.2 Alonso and Lamata's Approach

We have just seen Saaty's approach to determining whether a given matrix falls within our tolerances for inconsistency. This approach is very involved, and takes several steps after calculating the value of λ_{max} for the matrix in question. There is an easier way, developed by Alonso and Lamata [1], that allows us to immediately know upon calculation of λ_{max} whether a matrix falls within our tolerances. This method is presented here, and is the method we will use for the remainder of the paper when working with examples.

To begin, we first calculate the average value of λ_{max} for a large number of $n \times n$ matrices. We will refer to this value as $\overline{\lambda}_{max}$. In Alonso and Lamata's study, they used 500,000 randomly generated matrices for each n, with $3 \leq n \leq 13$. Note that these are the only values of n we are interested in for the AHP. For a matrix of size 2, we will always have a consistent matrix, and thus $\lambda_{max} = 2$. Furthermore, we split larger matrices into smaller, more manageable sizes (see Section 7.3 for a discussion of this process). By finding $\overline{\lambda}_{max}$ for the matrices of each size n, Alonso and Lamata were able to plot these values against the size of the matrix and find a least-square line to model the relationship. The curve that fits the data best is linear, with equation

$$\overline{\lambda}_{max}(n) = 2.7699n - 4.3513 \tag{10}$$

This line fits the data very well, with a correlation coefficient of 0.99 [1].

From this data, we can easily calculate the CR for any size of matrix. However, we will argue that it will be easier to simply use the value of λ_{max} generated by the ratio matrix in question. This way, we do not have to do any further calculation after we have found the weight matrix and λ_{max} in order to determine if the weight matrix is even valid. Hence, we do the following calculations to show the maximum acceptable value of λ_{max} if we want our CR to be less than 0.1.

First, we know that we can represent the RI for some matrix of size n by

$$RI = \frac{\overline{\lambda}_{max} - n}{n - 1}$$

Using the definitions of CI and CR (which we want to be less than 0.1), we can then see that

$$CR = \frac{\lambda_{max} - n}{\overline{\lambda}_{max} - n} < 0.1$$

Solving for λ_{max}, we have

$$\lambda_{max} < n + 0.1(\overline{\lambda}_{max} - n)$$

Thus, combining this result with our equation 10, we see that

$$\lambda_{max} < 1.17699n - 0.43513 \tag{11}$$

Using this equation, we can determine the maximum allowable λ_{max} for an $n \times n$ matrix by simply calculating $1.17699n - 0.43513$. If our λ_{max} is less than this value, then our matrix will still be within the allowable inconsistency, and we don't have to mess with the CI, CR, or RI at all. Furthermore, we can follow these same calculations to determine the allowable λ_{max} for other tolerances. If we are less concerned with the consistency of our data, and would rather not have to go back and reevaluate ratio matrices, we might pick a value larger than 0.1. For instance, if we were to allow a more inconsistent matrix, we could go up to a tolerance of 20%, which would correspond to a CR less than 0.2. Then, following through the same calculations here, we would come up with the equation

$$1.55398n - 0.87026$$

Given equation 11 we can create a table of the maximum allowable λ_{max} for a matrix of size n. From this table, we can easily determine if a given ratio matrix falls within our tolerances, or if we have to go back and try again. This table will be very useful as we go further in this paper. It is presented as Table 9.

Table 9: The maximum accepted λ_{max} for an $n \times n$ matrix [1]

n	λ_{max}	n	λ_{max}
3	3.0957	8	8.9806
4	4.2727	9	10.1576
5	5.4497	10	11.3346
6	6.6266	11	12.5166
7	7.8036	12	13.6886

Note that as we create larger and larger matrices, the largest allowable λ_{max} gets further and further from the n in question. This alludes to the idea that as we compare more and more alternatives, our results from the AHP are "less accurate." Methods for dealing with this problem are dealt with in the following section.

7.3 Clustering

(*note*: While working through this section, readers may find it helpful to refer to the worked example of clustering given in Section 7.3.1).

So far we have discussed AHP problems and examples with relatively few criteria and alternatives. In these cases, we can proceed just as we have been without getting too worried that the number of elements we are comparing could be contributing to error. As we look at more and more alternatives however, there certainly is more and more room for judgment error, leading to a less consistent result. Clustering allows us to look at the problem on a smaller scale which we are more used to, and then expand to the rest of the elements. Because we want to minimize inconsistency, whenever we are looking at more than 9 elements, be they alternatives or criteria, we will proceed by clustering our elements.

Consider the following example where we would have to use clustering [6]. We want to compare a bunch of apples, but there are so many of them that we are worried about inconsistency being too high. Hence, we set about clustering the apples into small groups of about seven apples each. We could, for example, cluster one way according to size, one way according to color, and another way according to age. In this manner, when we look at the alternatives in each cluster under the corresponding criteria, we are comparing apples that are close in respect to the common attribute. If there was large disparity in the alternatives, it would be more difficult to get an accurate comparison between two elements (see Figure 5). Hence, by clustering, we are making sure we are getting more accurate ratios between the elements.

When creating the clusters, it is important to have overlap of adjacent clusters. Consider our apple example, and the cluster of "size." In the cluster of the seven largest apples, we would then use the smallest apple in this cluster as the largest apple in the next cluster. For a visual representation of this clustering overlap, see Figure 6. In this manner, once we are done comparing within the ratios, we can find our weight matrix for all of the alternatives.

We first take the smallest element in the "largest" cluster (which we will call a, which in our apple example was the smallest "large" apple) and look at the cluster in which it is the largest element (i.e., the next "smaller" cluster). We then divide the weight of all the elements in this cluster by the weight of a as it is in this "smaller" cluster. Finally, to rejoin this cluster with the "larger" cluster, we multiply all of the weights by the weight of a in the "larger" cluster. We can then repeat this process with a new overlapping element, and working in this manner until we have gone through all of the clusters, we find the weights for all of the alternatives as a whole.

7.3.1 An example

To see just how clustering works, we will work through how it would work for the apples example mentioned earlier. Suppose we have 19 apples of various sizes, and we want to compare them under the criteria of size. We would then create three different clusters of 7 apples–namely small, medium, and large apples. We will discuss later just how we decide which apples belong to what cluster. For now, suppose we have made our clusters and gone through to find the weights of each alternative

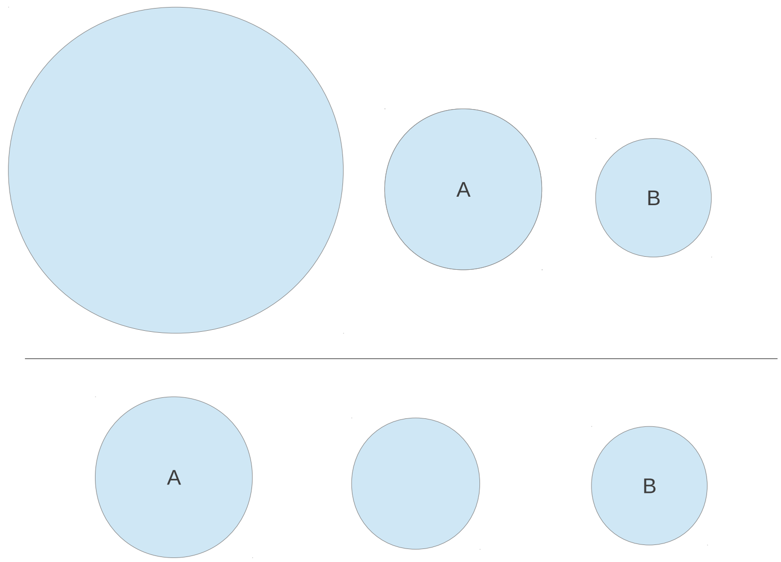


Figure 5: Circles A and B are the same size in each comparison, but when they are put with a circle closer to their size, it is easier to estimate just how much larger one is than the other [7]

just as we always have, but only focusing on one cluster at a time. Suppose we found the weights for the size of the apples as given in Table 10.

Note that the columns of this table are simply the normalized weight matrices that we are familiar with. The only difference is that they aren't the overall weights, but instead the weights for a small sample given by each cluster. We will eventually refine the weights in this table to get the final overall weights. Note also that the largest apple in the small cluster is also the smallest apple in the medium cluster, even though they have different weights. This is because when compared within either cluster, they will certainly have different weights because of the apples they are being compared to. Similarly, the largest medium apple is the same as the smallest large apple. We can then create a weight matrix for all of our alternatives by the following process. First, bring the medium cluster into conjunction with the large cluster by dividing by the largest medium apple's weight and then multiplying by the smallest large apple's weight (really, these two weights are the same since they come from the same apple). We then get a weight matrix for all of the apples in the medium cluster (given in 12) which now is a part of the large cluster, and so all the weights make sense within that cluster.

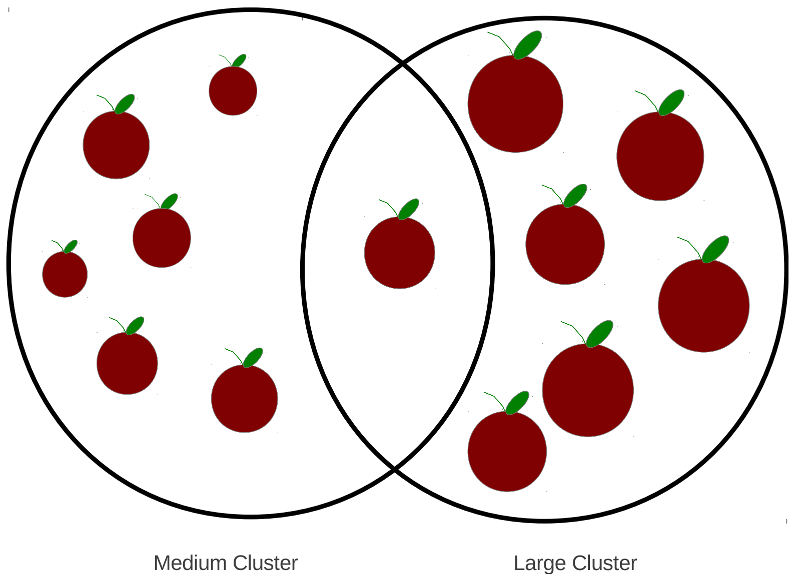


Figure 6: The structure of the “medium” and “large” clusters.

Table 10: The relative weights of apples under the criterion “size” by clustering

Small	Medium	Large
0.045	0.088	0.086
0.103	0.118	0.094
0.121	0.132	0.120
0.134	0.145	0.142
0.179	0.162	0.167
0.205	0.171	0.176
0.214	0.184	0.215

$$\frac{1}{0.184}\begin{bmatrix}0.088\\0.118\\0.132\\0.145\\0.162\\0.171\\0.184\end{bmatrix}0.086=\begin{bmatrix}0.041\\0.055\\0.062\\0.068\\0.076\\0.080\\0.086\end{bmatrix} \tag{12}$$

We then do the same thing for the small cluster, using our new value for the smallest "large" apple, 0.041.

$$\frac{1}{0.214}\begin{bmatrix}0.045\\0.103\\0.121\\0.134\\0.179\\0.205\\0.214\end{bmatrix}0.041=\begin{bmatrix}0.009\\0.020\\0.023\\0.026\\0.034\\0.039\\0.041\end{bmatrix}$$

Hence, combining our new found weight matrices with the matrix for the large cluster, we find that the weight matrix for size out of these 19 apples:

$$\begin{bmatrix}0.009\\0.020\\0.023\\0.026\\0.034\\0.039\\\color{red}{0.041}\\0.055\\0.062\\0.068\\0.076\\0.080\\\color{red}{0.086}\\0.094\\0.120\\0.142\\0.167\\0.176\\0.215\end{bmatrix}$$

which we would then normalize. Note that the entries in red are the "overlapping" alternatives.

We must now figure out just how we are going to create these clusters. In the example of apples of various sizes, it is easy to see how we could decide which apples belonged to which cluster, and could easily place them. However, with more abstract criteria and alternatives, the decision is not always as easy.

Saaty gives three possible approaches to determining clusters, ranked by least to most efficient [6]. The first of these is called the *elementary approach*. It is very simple and intuitive, but not very effective. It works by simple comparison. Given n alternatives, we can pick one out from the bunch. Then, working our way through the rest of the other alternatives, we compare pairwise with our selected alternative. If a compared alternative beats out the one we selected (if we find an apple bigger that the one in our hand, for example), we have a new "largest" alternative, and we continue comparing. Once we have found the largest, we would repeat the process with this alternative removed, trying to find the second largest. As can be seen, we would have to go through $(n-1)!$ comparisons just to figure out our clustering [6].

The second proposed method is called *trial and error clustering*. In this process, we try to get an idea of what the clusters would reasonably look like. We make three general groups, "large," "medium," and "small." We could look through the alternatives, see which looked to fit where ("that apple looks pretty small, we should probably put it in the small category"), and place them temporarily. We then place the alternatives in each group into several clusters, and make a pass through, making comparisons. Any misfits are moved to one of the other two groups. We then re-cluster, and make another comparison pass, again moving the misfits. After these two passes, the clusters should be reasonable enough that we can carry out our process as in the example.

The most efficient method is *clustering by absolute measurement*. In absolute measurement, the group looks at each alternative and tries to assign a weight without going through the whole AHP. This way, general clusters can be formed and then put through the AHP. The clusters are already comprised of elements that the group thinks go together, and the AHP simply refines the weight that they have assigned by using pairwise comparisons between the alternatives, rather than trying to assign some arbitrary number to each alternative.

8 Paradox in the AHP

We have now seen that there are matrices that are so inconsistent that they are weeded out before they even make it into our calculations to determine rankings. In this way, we discard matrices that could potentially give us rankings that aren't consistent with the ideas behind the AHP. However, what if we could create matrices that are within our tolerances for inconsistency, but, by their very structure, give paradoxical results? This section will examine some examples of these paradoxes, and will discuss the implications they have on the AHP.

8.1 Perturbing a Consistent Matrix

The first method we will examine for creating a paradox is that of simply making small alterations to a consistent matrix. In this example we will examine 5 alternatives, with weight matrix given in Table 11.

This weight matrix was created by assigning each alternative a weight from 2 to 6, and then normalizing the weight matrix. From this matrix, we could then create the consistent ratio matrix. However, in order to create a paradox, we change the preference ratio of $A : B$ in the ratio matrix. Hence, instead of $a_{12} = 0.4$ (which is simply 0.105/0.263, which we got from the weights of each alternative), we have $a_{12} = 5$, which we assign. We do the same swap with a_{21}, changing it to 0.2, or the inverse of 5. By doing this, we want to see if making one small change will go "unnoticed" when we calculate the new weight matrix. We then have the result shown in equation 13. Note that the changed entries are shown in red.

$$\begin{bmatrix} 1 & 5 & 0.5 & 1 & 0.333 \\ 0.2 & 1 & 1.25 & 2.5 & 0.833 \\ 2 & 0.8 & 1 & 2 & 0.666 \\ 1 & 0.4 & 0.5 & 1 & 0.333 \\ 3 & 1.2 & 1.5 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0.229 \\ 0.164 \\ 0.203 \\ 0.101 \\ 0.304 \end{bmatrix} = 5.909 \begin{bmatrix} 0.229 \\ 0.164 \\ 0.203 \\ 0.101 \\ 0.304 \end{bmatrix} \tag{13}$$

As we can see, our new weight matrix recognized the change we made, and adjusted accordingly. Now, since we decided we preferred A to B, even though the rest of the matrix is consistent with the case where we preferred B to A, the weight matrix tells us that A is a better choice than B. So, we can see that the process picked up on the change we made to just one entry. This one change does give us an eigenvalue that is outside of our acceptable range of inconsistency however, indicating how much changing just one entry from the consistent matrix can result in relatively large inconsistency (note also that since the original ratio matrix was consistent, it would have had a λ_{max} of 5 by Theorem 7).

We have now seen just how little it takes to upset a consistent matrix. Thus, we now try to create a matrix that falls within our tolerances for inconsistency. Let us try by changing the altered entry to 2 rather than 5, which indicates a less drastic reversal of preference. Hence, we have the new weight matrix for our alternatives given in Table 12.

Note that this weight matrix still retains the original preference of B over A as given in Table 11, and that the matrix falls within our tolerances for inconsistency.

Let us take a moment to consider what this situation tells us. When going through pairwise comparisons, we decided that alternative A was preferable to B. However, once we ran through the AHP, we are told that in actuality alternative B is preferable to alternative A. On one hand, we could argue that although alternative A is preferable to alternative B when we only consider those two alternatives, we are comparing A to B *only*. We are not taking into account any of the other alternatives, even though we are trying to decide which alternative is best out of a group of 5, not just a group consisting of A and B. While A may be preferable to B pairwise, the weight matrix tells us that alternative B is preferable to A in the larger scheme of things. This argument may be acceptable given certain situations. Consider the example of basketball teams playing one another given in Section 6. Although George Fox beat Whitworth, when we take into account all of the other teams, Whitworth actually did better in the conference.

This argument certainly makes sense for the given problem, but consider an entirely different situation. Now suppose we are back choosing a computer for the math lab given in Section 5.2, and we are choosing between 5 computers. Suppose we are evaluating based on the "looks" criterion, and so are trying to determine the prettiest computer. When comparing pairwise, we note that computer A looks better than computer B, but when we calculate the weight matrix, we find Table

Table 11: The weight matrix for our 5 alternatives

A	0.105
B	0.263
C	0.210
D	0.105
E	0.316

Table 12: The new weight matrix: $\lambda_{max} = 5.336$

A	0.166
B	0.201
C	0.211
D	0.105
E	0.316

12. Supposedly, given our pairwise rankings, B actually looks better than A. However, what if we start removing alternatives? If we take away all of the computers other than A and B, we see that the weight matrix tells us B is a more attractive option than A, which we know to be false. In this case, the paradox found does not makes sense given the problem.

This example illustrates several aspects of a paradox of this nature. First, we see that we can only perturb the matrix by a small amount. With a completely consistent matrix, we were only able to change one entry by a small amount before we were suddenly outside our tolerances for inconsistency. Moreover, we noticed that the paradoxical nature of the matrix is dependent on the actual problem we are looking at. This idea that the results must be interpret based on the original problem is crucial to the AHP. We cannot blindly follow our calculations without first considering what they are telling us.

8.2 Preference Cycles

The second paradox we will look at arises from the idea of *preference cycles*, which we define here.

Definition 16. *A* ***preference cycle*** *is a set of n elements with weights* $w_1, w_2, \ldots, w_n$ *such that*

$$w_1 > w_2 > \cdots > w_i > w_{i+1} > \cdots > w_n > w_1$$

A real world example of a preference cycle is the game of rock, paper, scissors, where each element beats exactly one other element, and loses to exactly one other element.

With this definition in mind, we will look to create a consistent weight matrix that arises from a situation where there is a preference cycle. Note that in the previous paradox example, a preference cycle arose when we changed the preferences of A and B. In this example, we will look what happens with a ratio matrix created solely from preference cycles. Hence, consider the ratio matrix for six alternatives given in Table 13.

Note that we have the preference cycle $A > B > C > D > E > F > A$, and that the matrix is within our inconsistency tolerances.

The case that we have presented here is interesting because of what the AHP tells us about our alternatives. When the weight matrix is calculated, we find that every alternative has the same weight. That is, we see that no one alternative is any better than another, which makes sense given the problem. Moreover, we can have alternatives that make no sense within a standard scale, but still pass our tolerances for inconsistency and give us some kind of result that is (by way of the AHP), in a standard scale. If we think back to our rock-paper-scissors example, the results shown above make sense. Given two pairings, we can obviously choose one alternative over the other, or else the two alternatives are equal (that is, neither player wins). However, before the game starts,

Table 13: The ratio matrix for a preference cycle of six elements- $\lambda_{max} = 6.5$

	A	B	C	D	E	F
A	1	2	1	1	1	$\frac{1}{2}$
B	$\frac{1}{2}$	1	2	1	1	1
C	1	$\frac{1}{2}$	1	2	1	1
D	1	1	$\frac{1}{2}$	1	2	1
E	1	1	1	$\frac{1}{2}$	1	2
F	2	1	1	1	$\frac{1}{2}$	1

we have no idea what the other player will show, and so every alternative is just as good as any. This is represented in the weight matrix by equivalent weights for each alternative, and in the ratio matrix by the parities entries. For the given example in Table 13, the 6×6 matrix is the first instance when we are within our inconsistency tolerances.

8.3 Further Paradoxes

These two examples are interesting in their own right, but there are many more possibilities for paradox within the AHP. Since it was not the goal of this paper to investigate paradoxes specifically, and since there is not a lot of available research on the types of problems illustrated here, we will stop our investigation of paradox here. We will however present some interesting extensions on the topic, which can be explored further.

First, much of the background and ideas for creating paradoxes in this paper arose from research into voting theory [2]. Thus, the idea of "strategic voting" (wherein an individual or small group can exploit the voting system to a desired end) is a natural extension of the idea of paradox within the AHP. Moreover, since the AHP is used often in business when meaningful, costly decisions are being made, it would be important to know whether the system can be subversively exploited by an individual or small group. As was illustrated in Section 8.1, it is possible to prefer one alternative over another, only to have the AHP tell you that actually, the preference should be the other way around.

Furthermore, it would be interesting to see which kinds of matrices are more susceptible to paradox. We have seen that larger matrices allow for more inconsistency (as indicated by Table 9), so we would think that these larger matrices would have more room for giving strange results. Another question to address would be to see if we could "spread out" the inconsistency more throughout the matrix, or if it is solely matrices that are almost perfectly consistent that allow for paradox.

One area of research that has been looked into, but which was not really discussed in this paper, is that of adding and subtracting alternatives and criteria. In an ideal process, simply adding or removing alternatives should do nothing to change the rank order of the alternatives, but this is not always the case. We alluded to this idea with the computer's "looks" in Section 8.1, but did not discuss at length.

9 An Applied AHP Problem

Now that we have a thorough understanding of the AHP, we will see just how it works in a real world decision problem. In this section, we look at the difficult decision of deciding where to attend graduate school given several choices. For this problem, we surveyed a senior at Whitman College who had been accepted to five schools for graduate chemistry programs. Thus, we define the problem as follows.

The goal: to choose the best graduate school for a particular individual (Tyler, our dauntless scholar).

9.1 The Criteria

These criteria were determined by speaking with Tyler, and figuring out what was important to him for choosing a grad school. Within each criterion, any sub-criteria are given in italics.

Location: Simply, where the school is located. This criteria has three sub-criteria making up what the "location" really entails. *Weather* is certainly an important factor to consider when picking a school. Just look at the difference between winters in the midwest and winters in southern California. The type of municipality the school is in also comes into play. *Rural and urban* schools offer different things when it comes to life outside of campus, and it is important to consider what kind of city or town you will be living in. Location can also mean how far away you are from the things you care about. *Distance* from your family and friends can be a good thing or a bad thing depending on how you look at it.

Financial Incentive: This criterion reflects the amount of money the school is willing to give, as well as the cost of living in the city where the school is located.

Ranking: Certainly ranking is important when choosing a school. The name Harvard carries much more weight than a local community college. This ranking can be looked at from two angles. First, the *prestige* associated with the name and school can be considered. Second, the numerical *rankings* published by various sources such as US News and World Report. Both of these factors carry some weight when trying to evaluate a school's rank.

Degree Program: It is important to consider what the degree program at each university actually entails. Some may have more limiting requirements, while others may require exciting hands-on research. Either way, it is an important factor to consider when making a decision about what school to go to.

Campus: The school's campus can be large or small, beautiful or ugly, and either way, it factors into the decision of what school to go to. The *grounds* of the school may be beautiful, with plenty of open fields, or they may be small and largely concrete. The *buildings* on campus are also important to consider. Is there a good library? Are the classrooms nice? Is there an on campus coffee shop where I can study? Finally, since we are dealing with a chemistry student, it is also important to factor in the *labs* when making a decision.

Faculty: The school itself may be great, but without excellent faculty, it isn't really worth choosing. Here we look at two factors, the student's *potential advisor*, and the *rest of the department*.

Vibe: This is a very important factor when choosing a school, but it is hard to quantify. The intangible feeling you get when walking around campus, and how you feel when you picture yourself going to a certain school are certainly very important in the final decision.

9.2 The Alternatives

For this decision problem, we considered the following five schools:

- The University of California, Berkeley. We will denote this alternative as simply "Berkeley."
- The University of Wisconsin, Madison, which we will denote "Madison."
- The University of Chicago, which we will denote "Chicago."
- The University of California, San Diego, which we will denote "UCSD."
- The University of Washington, Seattle, which we will denote "UW."

These schools comprise our alternatives. Hence, given the decision problem, we can lay out a hierarchy as illustrated in Figure 7.

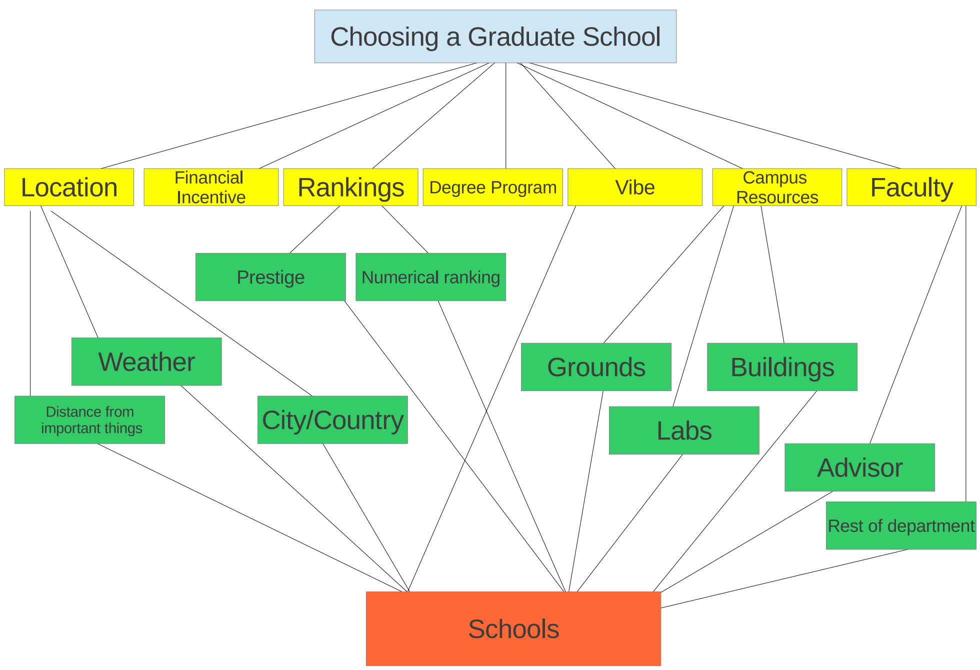


Figure 7: The hierarchy for our grad school problem

9.3 Gathering Data

After clearly defining the decision problem, the next step is to determine the ratio matrices for each criterion and alternative. We determined that the easiest way to do this would be through a worksheet of sorts, where Tyler could give scores for each comparison based on Saaty's fundamental scale (see Table 15). An example of one of these sheets is given in Table 16. Once all of this data was collected, we were able to create corresponding ratio and weight matrices for each criterion and alternative. These results are presented in Tables 17 through 52. Recall that when looking at these matrices, the preferences are read as row over column. That is, the number in entry a_{ij} represents how much alternative i is preferred over alternative j, according to the standard scale.

9.4 Determining the Final Weight Matrix

Just as we did in the computer example in Section 5.2, we simply multiply our way back up the hierarchy. For instance, tracing up "Madison" from its weight in the "Buildings" matrix, we first multiply the weight of "Madison" in this matrix by the weight of "Buildings" in its sub-criteria weight matrix. Then, this result is multiplied by the weight of "Campus" in its criteria matrix. We do this for every school, and for every path up the hierarchy. When we have all of these results, we add them up for each school. That is, all the weights for each school are summed separately, and these values become the final weight of the schools, giving us our result. The final weights from these calculations can be found in Table 14.

Table 14: Alternative Weight Matrix–"Overall"

School	**Weight**
Berkeley	0.4350
Madison	0.2172
Chicago	0.1743
UCSD	0.0706
UW	0.1060

As can be seen, Berkeley is the school that best satisfies our original goal of choosing the best grad school for Tyler, and by a generous margin. Before going through the process, Tyler stated that he was not sure where he was going to go. He knew that Berkeley and Madison were the two top contenders, and was interested to see if the process could set them apart from each other. With the results in Table 14, it is plain to see that the AHP is a powerful tool in making tough decisions.

Note that we can look at the "intermediate hierarchies" as well. If we were interested in what school best fulfilled the "campus resources" criterion, we could follow the same process as we did to calculate the values given in Table 14, but taking "campus" as the goal of our problem. Hence, the "campus resources" sub-criteria become the criteria in this problem. In this case, it turns out that Chicago best satisfies the goal. If Tyler was making his choice exclusively on how good the campus was, then Chicago would be the best bet. From this, we can see that Berkeley did not win out every criterion, and yet still won overall. This goes further to show how powerful a tool the AHP is. Even though an alternative does not win outright in every category, when everything is taken into account, we can find that the alternative is best given the entire problem.

9.5 Thoughts on the Applied Problem

By working through this problem, we have discovered just how straightforward and useful a tool the AHP is. A decision that was originally locked in stalemate was handily resolved by simple pairwise comparisons and some linear algebra. Moreover, we have seen just how easy it is to gather the data necessary to run through the process given just one person making the decision. If we were instead working through a problem that dealt input from multiple people, we would have to devise some way of fairly aggregating the responses of each individual. This aspect of the AHP was not investigated in this paper, nor was it discussed in Saaty's article [6]. It is certainly a very important aspect to decision problems, and should be considered when using the AHP to resolve group choices.

We have also shown (in our particular scenario) that it is not terribly difficult to arrive at weight matrices that are within our tolerances for inconsistency. Working through this problem, we had only one weight matrix out of 11 that was outside of our tolerance (the "urban/rural" alternative matrix). For this matrix, we went back and reevaluated the corresponding ratio matrix by running through the pairwise comparisons again, and quickly reached a weight matrix that fell within our tolerance.

10 Conclusions and Ideas for Further Research

As we have seen, the AHP is a useful, applicable tool to many situations in life. We are faced with decision problems every day, and often these are difficult to solve. With the AHP, we now have an excellent method for turning these difficult, incorporeal questions into reliable mathematics, from which we can easily determine a solution.

Furthermore, we have seen that this process is relatively easy. Solely through pairwise comparison and simple linear algebra, we can arrive at some powerful results. Moreover, these results actually mean something. That is, by running through the AHP we are calculating the ∞-dominance of the alternatives in question. Rather than a fancy trick of linear algebra, we actually have some theoretical backing to our results.

We have also seen that the process is resistant to strange results. By creating inconsistency tolerances, we get rid of many matrices that disagree with the principles at work in the AHP. However, this does not mean that the process is perfect, and completely foolproof. There are still possibilities for paradox and illogical results even with the inconsistency tolerances in place.

This is one very intriguing area of further research. While we have only presented two paradoxes in this paper, there are certainly countless more examples that can be dreamt up. A closer examination of some of these could prove to be very interesting indeed, as well as shedding light on just what can go wrong in the AHP. The specific idea of "strategic voting," a term borrowed from voting theory, would be very interesting. Since the AHP is used often in the business world to make important decisions, it would be worthwhile to investigate just how much one individual or a small group could strategically affect the overall rankings of the alternatives.

Additionally, further research into the kinds of problems the AHP is used to look at could be very interesting. Specifically, the area of sports problems (such as the basketball example given in Section 6) would be an attractive application. As far as I know, there has not been a lot of investigation as to using the AHP to look at these kinds of problems, but it seems like they are a prime situation for analysis by the process.

As I only conducted an applied problem for one person, gathering the data for the comparisons was relatively easy. It would be interesting to research different methods of gathering this data from a group of people, rather than an individual. This would certainly be a worthwhile investigation, since decisions are often made by groups rather than individuals, and we need an effective way of gathering and consolidating the opinions from everyone in the group. Moreover, it would be interesting to look at the differences between one of these group AHP problems compared to an individual AHP problem.

11 Appendix

Table 15: The Fundamental Scale [6]

Intensity of importance on an absolute scale	Definition	Explaination
1	Equal importance	Two activities contribute equally to the objective
3	Moderate importance of one over the other	Experience and judgment strongly favor one activity over another
5	Essential or strong importance	Experience and judgment strongly favor one activity over another
7	Very strong importance	An activity is strongly favored and its dominance demonstrated in practice
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values	When compromise is needed

Table 16: An example worksheet given to Tyler in order to gather data for the pairwise comparisons.

Instructions: Below are a list of criteria paired off with one another. In each pairwise comparison, circle the criteria which is more important in the decision of where to go to grad school. In the space provided, assign the intensity of the preference using the numerical values given in the Fundamental Scale. When making comparisons, only conisder the two criteria at a time, without regard for any other criteria.

First Criterion	**Second Criterion**	**Numerical Score**
Location	Financial Incentive	
Location	Ranking	
Location	Degree Program	
Location	Campus	
Location	Faculty	
Location	Vibe	
Financial Incentive	Ranking	
Financial Incentive	Degree Program	
Financial Incentive	Campus	
Financial Incentive	Faculty	
Financial Incentive	Vibe	
Ranking	Degree Program	
Ranking	Campus	
Ranking	Faculty	
Ranking	Vibe	
Degree Program	Campus	
Degree Program	Faculty	
Degree Program	Vibe	
Campus	Faculty	
Campus	Vibe	
Faculty	Vibe	

Table 17: Criteria Ratio Matrix

	Loc.	Fin.	Rank	Degree	Camp.	Fac.	Vibe
Loc.	1	5	6	6	4	$\frac{1}{3}$	$\frac{1}{4}$
Fin.	$\frac{1}{5}$	1	$\frac{1}{2}$	2	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{7}$
Rank	$\frac{1}{6}$	2	1	3	$\frac{1}{3}$	$\frac{1}{5}$	$\frac{1}{6}$
Degree	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$	1	$\frac{1}{3}$	$\frac{1}{5}$	$\frac{1}{7}$
Camp.	$\frac{1}{4}$	3	3	3	1	$\frac{1}{3}$	$\frac{1}{4}$
Fac.	3	4	5	5	3	1	$\frac{1}{5}$
Vibe	4	7	6	7	4	5	1

Table 18: Criteria Weight Matrix (Eigenvalue = 7.736)

Criterion	**Weight**
Loc.	0.1731
Fin.	0.0383
Rank	0.0490
Degree	0.0286
Camp.	0.0838
Fac.	0.2074
Vibe	0.4198

Table 19: Location Sub-Criteria Ratio Matrix

	Weather	Rural/Urban	Distance
Weather	1	4	$\frac{1}{3}$
Rural/Urban	$\frac{1}{4}$	1	$\frac{1}{6}$
Distance	3	6	1

Table 20: Location Sub-Criteria Weight Matrix (Eigenvalue = 3.054)

Sub-Criterion	**Weight**
Weather	0.2704
Rural/Urban	0.0852
Distance	0.6444

Table 21: Rankings Sub-Criteria Ratio Matrix

	Prestige	Rankings
Prestige	1	3
Rankings	$\frac{1}{3}$	1

Table 22: Rankings Sub-Criteria Weight Matrix (Eigenvalue = 2)

Sub-Criterion	**Weight**
Prestige	0.75
Rankings	0.25

Table 23: Campus Sub-Criteria Ratio Matrix

	Grounds	Buildings	Labs
Grounds	1	$\frac{1}{3}$	$\frac{1}{6}$
Buildings	3	1	$\frac{1}{5}$
Labs	6	5	1

Table 24: Campus Sub-Criteria Weight Matrix (Eigenvalue = 3.094)

Sub-Criterion	**Weight**
Grounds	0.0883
Buildings	0.1946
Labs	0.7171

Table 25: Faculty Sub-Criteria Ratio Matrix

	Potential Advisor	Rest of Department
Potential Advisor	1	$\frac{1}{4}$
Rest of Department	4	1

Table 26: Faculty Sub-Criteria Weight Matrix (Eigenvalue = 2)

Sub-Criterion	**Weight**
Potential Advisor	0.2
Rest of Department	0.8

Table 27: Alternative Ratio Matrix–“Weather”

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	6	7	$\frac{1}{3}$	4
Madison	$\frac{1}{7}$	1	2	$\frac{1}{7}$	2
Chicago	$\frac{1}{7}$	$\frac{1}{2}$	1	$\frac{1}{7}$	2
UCSD	3	7	7	1	8
UW	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$	1

Table 28: Alternative Weight Matrix–“Weather” (Eigenvalue = 5.283)

School	**Weight**
Berkeley	0.2928
Madison	0.0790
Chicago	0.0589
UCSD	0.5207
UW	0.0486

Table 29: Alternative Ratio Matrix–“Urban/Rural”

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	4	6	3	$\frac{1}{3}$
Madison	$\frac{1}{4}$	1	4	4	$\frac{1}{5}$
Chicago	$\frac{1}{6}$	$\frac{1}{4}$	1	$\frac{1}{4}$	$\frac{1}{6}$
UCSD	$\frac{1}{3}$	$\frac{1}{4}$	4	1	$\frac{1}{5}$
UW	3	5	6	5	1

Table 30: Alternative Weight Matrix–“Urban/Rural” (Eigenvalue = 5.436)

School	**Weight**
Berkeley	0.2725
Madison	0.1371
Chicago	0.0404
UCSD	0.0882
UW	0.4618

Table 31: Alternative Ratio Matrix–“Distance”

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	5	4	7	2
Madison	$\frac{1}{5}$	1	$\frac{1}{3}$	5	$\frac{1}{4}$
Chicago	$\frac{1}{4}$	3	1	5	$\frac{1}{4}$
UCSD	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{5}$	1	$\frac{1}{7}$
UW	$\frac{1}{2}$	4	4	7	1

Table 32: Alternative Weight Matrix–“Distance” (Eigenvalue = 5.384)

School	**Weight**
Berkeley	0.4255
Madison	0.0870
Chicago	0.1413
UCSD	0.0343
UW	0.3119

Table 33: Alternative Ratio Matrix–"Prestige"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	4	6	8	8
Madison	$\frac{1}{4}$	1	3	5	5
Chicago	$\frac{1}{6}$	$\frac{1}{3}$	1	3	3
UCSD	$\frac{1}{8}$	$\frac{1}{5}$	$\frac{1}{3}$	1	$\frac{1}{2}$
UW	$\frac{1}{8}$	$\frac{1}{5}$	$\frac{1}{3}$	2	1

Table 34: Alternative Weight Matrix–"Prestige" (Eigenvalue = 5.228)

School	**Weight**
Berkeley	0.5614
Madison	0.2289
Chicago	0.1104
UCSD	0.0428
UW	0.0565

Table 35: Alternative Ratio Matrix–"Ranking"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	4	5	6	8
Madison	$\frac{1}{4}$	1	3	4	6
Chicago	$\frac{1}{5}$	$\frac{1}{3}$	1	3	5
UCSD	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	1	3
UW	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{3}$	1

Table 36: Alternative Weight Matrix–"Ranking" (Eigenvalue = 5.330)

School	**Weight**
Berkeley	0.5293
Madison	0.2347
Chicago	0.1311
UCSD	0.0689
UW	0.0360

Table 37: Alternative Ratio Matrix–"Buildings"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	3	$\frac{1}{4}$	5	5
Madison	$\frac{1}{3}$	1	$\frac{1}{4}$	4	4
Chicago	4	4	1	6	5
UCSD	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{6}$	1	1
UW	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{5}$	1	1

Table 38: Alternative Weight Matrix–"Buildings" (Eigenvalue = 5.352)

School	**Weight**
Berkeley	0.2509
Madison	0.1452
Chicago	0.4977
UCSD	0.0517
UW	0.0546

Table 39: Alternative Ratio Matrix–"Grounds"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	6	4	4	3
Madison	$\frac{1}{6}$	1	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{3}$
Chicago	$\frac{1}{4}$	4	1	3	$\frac{1}{3}$
UCSD	$\frac{1}{4}$	3	$\frac{1}{3}$	1	$\frac{1}{4}$
UW	$\frac{1}{3}$	3	3	4	1

Table 40: Alternative Weight Matrix–"Grounds" (Eigenvalue = 5.402)

School	**Weight**
Berkeley	0.4585
Madison	0.0515
Chicago	0.1512
UCSD	0.0869
UW	0.2519

Table 41: Alternative Ratio Matrix–“Labs”

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	4	$\frac{1}{2}$	6	5
Madison	$\frac{1}{4}$	1	$\frac{1}{5}$	4	3
Chicago	2	5	1	6	5
UCSD	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$	1	$\frac{1}{3}$
UW	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{5}$	3	1

Table 42: Alternative Weight Matrix–“Labs” (Eigenvalue = 5.324)

School	**Weight**
Berkeley	0.3113
Madison	0.1278
Chicago	0.4347
UCSD	0.0417
UW	0.0733

Table 43: Alternative Ratio Matrix–“Potential Advisor”

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	1	2	6	5
Madison	1	1	2	6	5
Chicago	$\frac{1}{2}$	$\frac{1}{2}$	1	5	4
UCSD	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	1	$\frac{1}{2}$
UW	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{4}$	2	1

Table 44: Alternative Weight Matrix–“Potential Advisor” (Eigenvalue = 5.070)

School	**Weight**
Berkeley	0.3384
Madison	0.3384
Chicago	0.2078
UCSD	0.0465
UW	0.0687

Table 45: Alternative Ratio Matrix–"Rest of Department"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	4	5	6	8
Madison	$\frac{1}{4}$	1	3	5	6
Chicago	$\frac{1}{5}$	$\frac{1}{3}$	1	3	5
UCSD	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{3}$	1	3
UW	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{3}$	1

Table 46: Alternative Weight Matrix–"Rest of Department" (Eigenvalue = 5.353)

School	**Weight**
Berkeley	0.5269
Madison	0.2431
Chicago	0.1287
UCSD	0.0659
UW	0.0353

Table 47: Alternative Ratio Matrix–"Financial Incentive"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	$\frac{1}{3}$	$\frac{1}{6}$	3	$\frac{1}{4}$
Madison	3	1	$\frac{1}{3}$	4	$\frac{1}{3}$
Chicago	6	3	1	7	5
UCSD	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{7}$	1	$\frac{1}{6}$
UW	4	3	$\frac{1}{5}$	6	1

Table 48: Alternative Weight Matrix–"Financial Incentive" (Eigenvalue = 5.410)

School	**Weight**
Berkeley	0.0697
Madison	0.1394
Chicago	0.5174
UCSD	0.0390
UW	0.2345

Table 49: Alternative Ratio Matrix–"Degree Program"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	4	5	6	6
Madison	$\frac{1}{4}$	1	3	4	5
Chicago	$\frac{1}{5}$	$\frac{1}{3}$	1	3	4
UCSD	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	1	2
UW	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{2}$	1

Table 50: Alternative Weight Matrix–"Degree Program" (Eigenvalue = 5.320)

School	**Weight**
Berkeley	0.5231
Madison	0.2349
Chicago	0.1305
UCSD	0.0652
UW	0.0463

Table 51: Alternative Ratio Matrix–"Vibe"

	Berkeley	Madison	Chicago	UCSD	UW
Berkeley	1	3	4	6	5
Madison	$\frac{1}{3}$	1	3	5	4
Chicago	$\frac{1}{4}$	$\frac{1}{3}$	1	4	3
UCSD	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	1	$\frac{1}{3}$
UW	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	3	1

Table 52: Alternative Weight Matrix–"Vibe" (Eigenvalue = 5.314)

School	**Weight**
Berkeley	0.4704
Madison	0.2620
Chicago	0.1440
UCSD	0.0446
UW	0.0791

References

[1] Jose Alonso and Teresa Lamata, *Consistency in the Analytic Hierarchy Process: A New Approach.* International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. Vol. 14, No. 4 (2006) 445-459.

[2] Steven J. Brams, *Mathematics and Democracy: Designing better voting and fair-division procedures.* Princeton University Press. (2008).

[3] Godfrey Harold Hardy, *Divergent Series.* Clarendon Press, Oxford. (1956).

[4] Roger A. Horn, Charles R. Johnson, *Matrix Analysis.* Cambridge University Press, Cambridge. (1985).

[5] David C. Lay, *Linear Algebra and its Applications, Third Edition* Addison Wesley, 2006. Page 114.

[6] Thomas Saaty, *How to make a decision: The Analytic Hierarchy Process.* European Journal of Operational Research 48 (1990) 9-26. North Holland

[7] Xiaojun Yang, Liaoliao Yan, and Luan Zeng *How to handle uncertainties in AHP: The Cloud Delphi hierarchical analysis.* Journal of Information Sciences 222 (2013) 384-404.