A (virtual) tour of functional archetypal analysis

Irene Epifanio López
Dpt. Matemàtiques - IF, Univ. Jaume I (SPAIN)
epifanio@uji.es; http://www3.uji.es/~epifanio
Outline

- Archetypal analysis (AA)
- AA and ADA for multivariate data
  - Some applications: Industrial design, etc.
- AA and ADA for functional data
  - Application: Human development around the world over the last 50 years
- Other derived methodologies
- More applications
- Future work or working now
Archetypes

- Archetype (Wikipedia): from Greek, ἀρχή, archē, "beginning, origin", and τύπος, tupos, "pattern," "model," or "type"; original pattern from which copies are made.
Archetypes in real world
Archetype analysis (AA)

- Archetype concept in Statistics is the same as in life common.
- Objective (Cutler and Breiman, 1994): to find a few, not necessarily observed, extremal cases or pure types (the archetypes) such that:

1. all the observations are approximated by convex combinations of the archetypes, and
2. all the archetypes are convex combinations of the observations.
Archetypes in 2D

2 Prototypes

4 Prototypes

8 Prototypes

AA

K-means
Archetypes in 2D

Figure 2: Visualization of three archetypes.
Objective (Vinue, Epifanio, Alemany, 2015): to find a few, observed, extremal cases or pure types (the archetypoids) such that:

1. all the observations are approximated by convex combinations of the archetypoids, and

2. all the archetypoids are real observations.
Interpretation

ADA

K-means
Let $X$ be an $n \times m$ matrix with $n$ observations and $m$ variables. The objective of AA is to find the matrix $Z$ of $k$ $m$-dimensional archetypes. AA computes two matrices $\alpha$ and $\beta$ which minimize the residual sum of squares (RSS):

$$RSS = \sum_{i=1}^{n} \left\| x_i - \sum_{j=1}^{k} \alpha_{ij} z_j \right\|^2 = \sum_{i=1}^{n} \left\| x_i - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl} x_l \right\|^2,$$

under the constraints

1) $\sum_{j=1}^{k} \alpha_{ij} = 1$ with $\alpha_{ij} \geq 0$ for $i = 1, \ldots, n$ and

2) $\sum_{l=1}^{n} \beta_{jl} = 1$ with $\beta_{jl} \geq 0$ for $j = 1, \ldots, k$. 
ADA for multivariate data

The objective of ADA is to find the matrix $Z$ of $k \times m$-dimensional archetypoids (real cases). ADA computes two matrices $\alpha$ and $\beta$ which minimize the residual sum of squares (RSS):

$$RSS = \sum_{i=1}^{n} \left\| x_i - \sum_{j=1}^{k} \alpha_{ij} z_j \right\|^2 = \sum_{i=1}^{n} \left\| x_i - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl} x_l \right\|^2,$$

under the constraints

1) $\sum_{j=1}^{k} \alpha_{ij} = 1$ with $\alpha_{ij} \geq 0$ for $i = 1, \ldots, n$ and

2) $\sum_{l=1}^{n} \beta_{jl} = 1$ with $\beta_{jl} \in \{0, 1\}$ and $j = 1, \ldots, k$. 
Table 1: Relationship between archetypoid analysis and several unsupervised methods, as in Mørup and Hansen (2012): Principal component analysis (PCA), Non-negative matrix factorization (NMF), Convex NMF (CNMF), Archetype analysis (AA), Archetypoid analysis (ADA), Soft $k$-means (i.e. fuzzy $k$-means or the EM-algorithm for clustering), $k$-means and $k$-medoids. $\mathbb{B}$ represents the set \{0, 1\}.

| Method          | \( \beta \in \mathbb{R} \) | \( \alpha \in \mathbb{R} \) | \( X'\beta \geq 0 \) | \( \alpha \geq 0 \) | \( \beta \geq 0 \) | \( \alpha \geq 0 \) | \( |\beta_k|_1 = 1, \beta \geq 0 \) | \( |\alpha_n|_1 = 1, \alpha \geq 0 \) | \( |\beta_k|_1 = 1, \beta \in \mathbb{B} \) | \( |\alpha_n|_1 = 1, \alpha \geq 0 \) | \( \beta_{k,n} = \frac{\sum_{\bar{n}} \alpha_{k,\bar{n}}}{|\alpha_n|_1} \) | \( |\alpha_n|_1 = 1, \alpha \geq 0 \) | \( |\beta_k|_1 = 1, \beta \geq 0 \) | \( |\alpha_n|_1 = 1, \alpha \in \mathbb{B} \) | \( |\beta_k|_1 = 1, \beta \in \mathbb{B} \) | \( |\alpha_n|_1 = 1, \alpha \in \mathbb{B} \) |
|-----------------|------------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|-------------------|-----------------|-----------------|------------------|------------------|------------------|-----------------|------------------|------------------|-----------------|-----------------|-----------------|
| PCA             | \( \beta \in \mathbb{R} \) | \( \alpha \in \mathbb{R} \) |                 |                  |                 |                  |                 |                  |                  |                 |                 |
| NMF             | \( X'\beta \geq 0 \) |                  | \( \alpha \geq 0 \) | \( \alpha \geq 0 \) | \( \beta \geq 0 \) | \( \alpha \geq 0 \) | \( |\beta_k|_1 = 1, \beta \geq 0 \) | \( |\alpha_n|_1 = 1, \alpha \geq 0 \) | \( |\beta_k|_1 = 1, \beta \in \mathbb{B} \) | \( |\alpha_n|_1 = 1, \alpha \geq 0 \) | \( \beta_{k,n} = \frac{\sum_{\bar{n}} \alpha_{k,\bar{n}}}{|\alpha_n|_1} \) | \( |\alpha_n|_1 = 1, \alpha \geq 0 \) | \( |\beta_k|_1 = 1, \beta \geq 0 \) | \( |\alpha_n|_1 = 1, \alpha \in \mathbb{B} \) | \( |\beta_k|_1 = 1, \beta \in \mathbb{B} \) | \( |\alpha_n|_1 = 1, \alpha \in \mathbb{B} \) |
AA solution

- Cutler and Breiman (1994) proposed an alternating minimizing algorithm.
- Implemented in R by Eugster and Leisch (2009):
  
  ![R package archetypes](https://example.com/archetypes)  

- To solve the convex least squares problems, they used a penalized version of the non-negative least squares algorithm by Lawson and Hanson (1974).
Other AA solutions

- More efficient alternative algorithms for computing AA have been proposed, especially for large data sets, such as the implementation by
  - Chen et al. (2014).
  - Mair et al. (2017).
ADA solution

- It consists of two phases, a BUILD step and a SWAP step.
  - An initial set of archetypoids is computed in the BUILD phase.
  - The SWAP step seeks to improve the set of archetypoids by exchanging chosen observations for unselected cases and by checking if these replacements reduce the RSS.
- Implemented in R: package Anthropometry, adamethods (for big data sets: subsampling, Vinué, Epifanio (2020)).
Comparison

Figure 4: Archetypes (with red crosses) and archetypoids (with solid black circles) for simulated Bivariate Normal Data, with $k = 4$, together with the four representatives (with blue squares) obtained for the following methods, respectively: (a) SiVM, (b) SMRS, (c) AP, (d) HOTTOPIXX, (e) BPM and (f) classical clustering algorithms (PAM with blue squares, $k$-means with green triangles and fuzzy $k$-means with magenta diamonds).
Location

- \( k = 1 \): the archetype is equal to the mean and to the medoid in case of the archetypoid (Kaufman and Rousseeuw, 1990).

- \( k > 1 \): archetypes are on the boundary of the convex hull of the data (Cutler and Breiman, 1994), although this does not necessarily happen for archetypoids (Vinue et al., 2015).
Which k???

- Elbow criterion
Example of AA application: Design of USAF aircraft cockpits

1. Thumb-Tip Reach (Alcance de la punta del dedo pulgar)
2. Buttock-Knee Length, (Longitud nalga-rodilla)
3. Popliteal Height-Sitting (Altura de la rodilla, sentado)
4. Sitting Ht (Altura, hasta la cabeza, sentado)
5. 5. Eye Ht-Sitting (Altura de los ojos sentado)
6. 6. Acromion Ht – Sitting (Altura del hombro sentado)

Figure 1: Generic skeleton for an aircraft pilot.
Bad designs → problems

- Control Authority
- Eyes below glare shield
- External Vision
- Glare shield line projected to pilot's head
- Missed rudder throw by several inches
- Missed reach to full throttle by several inches
- Leg Clearance
- Missed forward stick by several inches
- Body Clearances
- Reach to Controls
Calculation of archetypes with 95% accommodation

Objective: To find cases on the border, i.e., extreme cases, for covering certain percentage of users. If the design enables extreme cases to operate efficiently, all other less extreme body types and size will be well accommodated.

How to find the archetypes so that the design fits 95% of the population?

1. Standardization.
2. Removing the more extreme 5% data (use of Chi-Square distribution with Mahalanobis distance or depth techniques).
3. AA application
Seven archetypes

1. High percentiles in all variables
2. Low percentiles in all variables
3. High percentiles in extremities and middle in torso
4. High in torso and medium in extremities
5. Medium-high for extremities, low torso
6. High percentiles in all variables, except arm
7. Medium-high for torso, low limbs

Table 5: Percentile values for seven archetypes

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thumb Tip Reach</td>
<td>94</td>
<td>2</td>
<td>99</td>
<td>44</td>
<td>68</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Buttock-Knee Length</td>
<td>97</td>
<td>1</td>
<td>86</td>
<td>7</td>
<td>89</td>
<td>66</td>
<td>22</td>
</tr>
<tr>
<td>Popliteal Height Sitting</td>
<td>94</td>
<td>2</td>
<td>96</td>
<td>37</td>
<td>77</td>
<td>87</td>
<td>0</td>
</tr>
<tr>
<td>Sitting Height</td>
<td>99</td>
<td>0</td>
<td>51</td>
<td>86</td>
<td>3</td>
<td>86</td>
<td>57</td>
</tr>
<tr>
<td>Eye Height Sitting</td>
<td>99</td>
<td>1</td>
<td>64</td>
<td>85</td>
<td>2</td>
<td>84</td>
<td>64</td>
</tr>
<tr>
<td>Shoulder Height Sitting</td>
<td>99</td>
<td>1</td>
<td>29</td>
<td>93</td>
<td>13</td>
<td>68</td>
<td>57</td>
</tr>
</tbody>
</table>
Seven archetypes

1. High percentiles in all variables
2. Low percentiles in all variables
3. High percentiles in extremities and middle in torso
4. High in torso and medium in extremities
5. Medium-high for extremities, low torso
6. High percentiles in all variables, except arm
7. Medium-high for torso, low limbs
Design of USAF aircraft cockpits

Once the extreme cases are clear, the cockpit can be designed.
Example ADA: Sport analytics (toy examples)

Fig. 1 cand$_b$ players (with red crosses, obtained by Eugster (2012)) and archetypoid players (with solid black circles and frame box) for the total minutes played and field goals made by a set of NBA players from the 2009/2010 season, together with the representatives (with blue squares) obtained for the following methods, respectively: (a) SVM and SMRS (not indicated because they match the cand$_b$ players), (b) AP, (c) BPM and (d) PAM, k-means and fuzzy k-means (blue squares, green triangles and magenta diamonds, respectively). The RSS are: 0.00165 (ADA) and 0.00169 (cand$_b$, SVM and SMRS). The computational times are: AA 2 sec.; ADA for each initial candidate set 25 sec.; SVM 〈 0.1 sec.; SMRS, 8 sec. (for regularization parameter 20: for others, for example 14 sec.).

Fig. 8 Star plot of alpha values for the 5 archetypoid teams (RAY in black, MAL in red, BAR in green, COR in blue and RSC in yellow) in the 2014-2015 Spanish football league. The final league classification appears in brackets.
Example of AA application: Texture segmentation

Fig. 11. Unsupervised segmentation of orthophotos using our procedure.

Fig. 12. Unsupervised segmentation of orthophotos by clustering the local granulometries.
Objective of functional AA (FAA): to find $k$ archetype functions (mixture of the data),

Objective of functional ADA (FADA): to find $k$ functions of the sample (archetypoids),

in such a way that our functional data sample can be approximated by mixtures of those archetypal functions.

- The vector norms are replaced by functional norms ($L^2$ norm, $\|f\|^2 = <f, f> = \int_a^b f(t)^2 dt$) in equation 1 and 2; the vectors $x_i$ and $z_j$ correspond to the functions $x_i$ and $z_j$.
- The meaning of $\alpha$ and $\beta$ in the functional case is identical to the multivariate case.
Computational details: basis approach

- Each function $x_i$ is expressed as a linear combination of known basis functions $B_h$ with $h = 1, ..., m$:

$$x_i(t) = \sum_{h=1}^{m} b_i^h B_h(t) = b'_i B$$

- $b_i$ : the vector of the coefficients
- B the functional vector whose elements are the basis functions.
Computational details: basis approach

\[ RSS = \sum_{i=1}^{n} \|x_i - \sum_{j=1}^{k} \alpha_{ij}z_j\|^2 = \sum_{i=1}^{n} \|x_i - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl}x_l\|^2 = \]

\[ \sum_{i=1}^{n} \|b'_iB - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl}b'_lB\|^2 = \sum_{i=1}^{n} \|(b'_i - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl}b'_l)B\|^2 = (3) \]

\[ \sum_{i=1}^{n} \|a'_iB\|^2 = \sum_{i=1}^{n} <a'_iB, a'_iB> = \sum_{i=1}^{n} a'_iW a_i, \]

where: \( a'_i = b'_i - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl}b'_l \)

and \( W \) is the matrix containing the inner products of the pairs of basis functions.

\[ w_{m_1, m_2} = \int B_{m_1}B_{m_2} \]

Constraints for \( \alpha \) and \( \beta \) identical as the multivariate case.
In the case of an orthonormal basis such as Fourier, $W$ is the order $m$ identity matrix, and $FAA$ ($FADA$, respectively) is reduced to $AA$ ($ADA$, respectively) of the basis coefficients.

But, in other cases, we may have to resort to numerical integration to evaluate $W$, but once $W$ is computed, no more numerical integrations are necessary.
Multivariate FAA and FADA

- Key: to define an inner product between multivariate functions, which is computed simply as the sum of the inner products of the components.
- FAA or FADA for $M$ multivariate functions is equivalent to $M$ independent FAA or FADA, respectively, with shared parameters $\alpha$ and $\beta$. 
Multivariate FAA and FADA computation

- Let \( f_i(t) = (x_i(t), y_i(t)) \) be a bivariate function. Its squared norm:
  \[
  \|f_i\|^2 = \int_a^b x_i(t)^2\,dt + \int_a^b y_i(t)^2\,dt
  \]

- The coefficients for \( x_i \) and \( y_i \) respectively for the basis functions \( B_h \) are \( b^x_i \) and \( b^y_i \)

\[
RSS = \sum_{i=1}^n \| f_i - \sum_{j=1}^k \alpha_{ij} z_j \|^2 = \sum_{i=1}^n \| f_i - \sum_{j=1}^k \alpha_{ij} \sum_{l=1}^n \beta_{jl} x_l \|^2 =
\]

\[
\sum_{i=1}^n \| x_i - \sum_{j=1}^k \alpha_{ij} \sum_{l=1}^n \beta_{jl} x_l \|^2 + \sum_{i=1}^n \| y_i - \sum_{j=1}^k \alpha_{ij} \sum_{l=1}^n \beta_{jl} y_l \|^2 =
\]

\[
\sum_{i=1}^n a^{x'}_i W a^x_i + \sum_{i=1}^n a^{y'}_i W a^y_i,
\]

where

\[
a^{x'}_i = b^{x'}_i - \sum_{j=1}^k \alpha_{ij} \sum_{l=1}^n \beta_{jl} b^{x'}_l \quad \text{and} \quad a^{y'}_i = b^{y'}_i - \sum_{j=1}^k \alpha_{ij} \sum_{l=1}^n \beta_{jl} b^{y'}_l
\]
Two indicators of World Bank Open Data:

- **Total fertility rate (TFR):** no. children that would be born to a woman if she were to live to the end of her childbearing years and bear children in accordance with current age-specific fertility rates.

- **Life expectancy at birth (LEB):** the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.

The series of each country goes from 1960 to 2013.
Application

- **190 countries considered.**
- All functions (those with or without missing years) are expressed with **32 B-spline basis functions of order 4 (cubic splines)** from 1960 to 2013, with equally spaced knots.
- TFR and LEB are measured in non-compatible units, so each functional variable should be standardized.
- Bivariate FADA with $k = 5$ archetypoids.
5 functional archetypoids
Leshoto

- TFR has decreased from nearly 6 children in 1960 to 3.
- LEB curve reflects a significant problem in Southern Africa: HIV/AIDS.
Channel Islands

- Representative of countries with low TFR and high LEB over the years.
Niger

- Representative of countries with high TFR over the years, but low LEB (36 years) in the 1960s, which has increased to nearly 60 years nowadays.
Qatar and Bhutan

- TFR has decreased spectacularly, from nearly 7 in the 1960s to 2 nowadays. But, this decrease has taken place at different times.
- LEB has increased considerably.
Countries with indicator curves similar to Lesotho are their neighboring countries, which are the countries most affected by HIV/AIDS.
The countries whose indicator functions coincide with those of the Channel Islands are Japan, Australia, North America and most European countries, and to a lesser extent, countries such as Russia and Argentina.
Abundance map for Niger

- Countries which mainly share their indicator functions are those in **Central Africa and Afghanistan**.
Abundance map for Qatar

- Countries with a similar behavior: majority of countries in the Arabian peninsula and neighboring countries, many countries in Central America and several in South America, several in Asia and countries in North Africa, although, those North African countries also share characteristics with Bhutan.
- Morocco, Algeria and Tunisia are a mixture of Qatar and Bhutan. Other countries are also a mixture of two or three profiles. For example, Turkey is a mixture between 30% the Channel Islands, 20% Qatar and 50% Bhutan.
Robust AA and ADA and FAA and FADA

- The RSS is formulated as the sum of the squared (vectorial or functional) norm of the residuals, $r_i$ ($i = 1, ..., n$).
- Here, $r_i$ denote vectors of length $m$ in the multivariate case or (univariate or multivariate) functions in the functional case. The least squares loss function does not provide robust solutions. So, we use Tukey weight or bisquare family of loss function.

Therefore, RSS in equations 1, 2, 3, 4 are replaced by $\sum_{i=1}^{n} \rho_c(||r_i||)$, where $|| \cdot ||$ denotes the Euclidean norm for vectors, the $L^2$-norm for univariate functions and corresponding norm for $P$-variate functions, and $\rho_c(||r_i||)$ is given by

$$\rho_c(||r_i||) = \begin{cases} 
\frac{c^2}{6} \cdot (1 - (1 - ||r_i||^2/c^2)^3) & \text{if } 0 \leq ||r_i|| \leq c \\
\frac{c^2}{6} & \text{if } c < ||r_i||
\end{cases}$$
Detecting outliers with AA and ADA

- Vinué, Epifanio (2020): the norm of the residuals of robust FADA are used to detect functional outliers, with univariate and multivariate functions.

- Cabero, Epifanio et al. (2021): Archetypes lie on the boundary of the convex hull of the data, meaning that they are extreme profiles. This makes AA sensitive to outliers, and we will take advantage of this in order to detect outliers.

  - Apply k-NN (sum of distance to k nearest neighbors) for a certain k for the \( \alpha \) matrices from elbow to P. Then, the \( P - e + 1 \) outlier scores obtained in each subspace are merged by a cumulative-sum approach, which is equivalent to averaging the scores.
Detecting outliers with AA and ADA

- Cabero, Epifanio et al. (2021): Archetypes lie on the boundary of the convex hull of the data, meaning that they are extreme profiles. This makes AA sensitive to outliers, and we will take advantage of this in order to detect outliers.
  - Apply k-NN (sum of distance to k nearest neighbors) for a certain k for the $\alpha$ matrices from elbow to P. Then, the P - e + 1 outlier scores obtained in each subspace are merged by a cumulative-sum approach, which is equivalent to averaging the scores.

**Fig. 1.** Example 1: (a) plot of the data set (see the text for details); (b) Screeplot; (c) Ternary plot.
AA and ADA with other kinds of data

- **Landmarks (Epifanio et al. (2018))**: As shape space is not a vectorial space, we work in the tangent space, the linearized space about the mean shape.

- **Missing data (Epifanio et al (2020))**:
  - **Mørup, Hansen (2012)**:
    \[
    \min_{s,c} \sum_{n,m} q_{n,m} \left( x_{n,m} - \sum_d \frac{\sum_m x_{n,m} c_{m,d}}{\sum_m q_{n,m} c_{m,d} + \epsilon} s_{d,m} \right)^2
    \]
    where \( Q \) is an indicator matrix such that \( q_{n,m} = 1 \) if the \( n \)th movie was rated by the \( m \)th user and zero otherwise and \( \epsilon = 10^{-3} \) is a small constant used for numeric stability and to avoid potential

- **Use of AA with missing data as collaborative filtering technique (Alcacer, Epifanio et al. 2021)**.

- **Mørup, Hansen (2012)**: They also proposed kernel-AA and relaxation of AA model (\( \beta > 1 \)).
AA and ADA with other kinds of data

- **Binary data** (Cabero, Epifanio (2020)): ADA works better than PAA and AA.

  Probabilistic AA (PAA) (Seth, Eugster, (2016)): The idea underlying PAA is to work in a parameter space instead of the observation space, since the parameter space is often vectorial even if the sample space is not.

- **Ordinal data** (Fernández, Epifanio, McMillan, 2nd revision).
More applications

- Biology (D’Esposito et al (2012)).
- Developmental psychology (Ragozini et al (2017)).
- Didactics (Cabero, Epifanio (2020)).
- E-learning (Theodosiou et al (2013)).
- Engineering (Millán-Roures, Epifanio et al. (2018)).
- Finance (Molinier, Epifanio (2019)).
- Genetics (Thøgersen et al (2013)).
- Machine learning problems: text mining, collaborative filtering, chemistry, computer vision (Mørup, Hansen (2012)).
- Market research (Li et al (2003); Porzio et al (2008)).
- Neuroscience (Tsanousa et al (2015); Hinrich et al (2016)).
Future work or working now

- Biarchetype analysis (Alcacer, Epifanio (2020)): It allows the extraction of extreme cases from both the observations and the variables.

\[
RSS = \sum_{i=1}^{n} \sum_{j=1}^{m} \left| x_{ij} - \sum_{g=1}^{k} \sum_{h=1}^{c} \alpha_{ijg} \left( \sum_{l=1}^{n} \sum_{r=1}^{m} \beta_{glr} x_{lr} \theta_{rh} \right) \gamma_{hj} \right|^2
\]

subject to the following restrictions

1. \( \sum_{g=1}^{k} \alpha_{ijg} = 1 \) con \( \alpha_{ijg} \geq 0 \) para \( i = 1, \ldots, n \).
2. \( \sum_{h=1}^{c} \gamma_{hj} = 1 \) con \( \gamma_{hj} \geq 0 \) para \( j = 1, \ldots, m \).
3. \( \sum_{l=1}^{n} \beta_{glr} = 1 \) con \( \beta_{glr} \geq 0 \) para \( g = 1, \ldots, k \).
4. \( \sum_{r=1}^{m} \theta_{rh} = 1 \) con \( \theta_{rh} \geq 0 \) para \( h = 1, \ldots, c \).

In matrix notation, the archetypes in \( Z \) are

\[
Z_{k \times c} = \beta_{k \times n} X_{n \times m} \theta_{m \times c}
\]

and, in turn, data are approximated by

\[
X_{n \times m} \simeq \alpha_{n \times k} Z_{k \times c} \gamma_{c \times m}
\]

- Other kind of data: mixed, censored, directional, text data, ...; Regularized AA.

- Open questions: everything done with clustering can be moved to AA.
References


References


THANKS FOR YOUR ATTENTION!

http://www3.uji.es/~epifanio