



Visualizing the Phylomemetic Tree Innovation as Evolutionary Process

DENI KHANAFIAH
Dept. Computational Sociology
Bandung Fe Institute
dk@students.bandungfe.net

HOKKY SITUNGKIR
Dept. Computational Sociology
Bandung Fe Institute
hs@compsoc.bandungfe.net

Abstract

The Innovation of Artifacts is somewhat can be seen as a process of evolution. The paper presents an endeavor to view the evolution of artifact by using evolutionary concept of memetics. We showed the ways to build a phylomemetic tree based on memes constituting an artifact to infer or estimate the evolutionary history and relationship between artifacts. UPGMA algorithm and the Shortest Tree Method using Minimum Spanning Tree (MST) techniques are presented to construct the phylomemetic tree of innovation. To show an implementation, we use innovation of cellphone as an example.

Keywords: artifact, innovation, evolution, memetic, phylomemetic tree

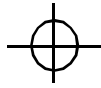
1. Innovation in the Evolution of Technological Artifacts

Innovation is one of many processes that may emerge new artifacts. As a (complex) system, innovation of artifact means the change of states component of the system, in so forth emerging a system, which its characters or behaviors are different from the previous one (Frenken, 2001a). A process of innovation can be regarded as relatively random process (Mokyr, 1997; Frenken, 2001a; Kauffman, 1995), in which means the resulting technology cannot always be known precisely to fit with the environment or not.

Evolution process from technological artifacts can be viewed as the phenomenon of the emergence of a new type of artifacts by means of innovation - in turns replacing the old one. Up to these days, studies or analyses to understand the innovation in evolution or development of things over time are still an interesting field, especially to obtain deeper understanding of the principles behind its evolution (Kaplan et al, 2003).

Technological artifact changes from time to time. Changes may happen over the design and technology. In this paper, we try to see the possibility of using memetic in order to understand the evolution of technology artifacts. Memetic itself is the theory that described the social evolution, including cultural evolution, by using Darwinian paradigm, i.e. evolution through natural selection. Memetics becomes an alternative analytical tool to understand socio-culture phenomena (Situngkir,





2004), including innovation of technological artifacts.

One of analyses we do in understanding artifact evolutionary process with memetics is analysis of evolutionary relationship (ancestral relationship of artifact). In this analysis, we adopt the phylogenetic concept - a concept used in biology to infer the history and the evolutionary relationship of organisms based on their relativity characters. This concept had also considered as potential tool to analyze the change and evolution of organization and management process (Joel, *et al*, 2002). Using the same paradigm, we infer or estimate the evolutionary relationship of artifacts based on memes in artifact as their distinguished properties. The evolutionary relationship among different types of artifact is represented as branches, a tree-like diagram we defined as phylomemetic tree.

2. Innovation as a Memetic Evolution

As explained by Sartika (2004), memetics has many definitions since the term and the concept coined by Dawkins (1976). Some definitions include meme as a replicator unit (Blackmore, 1998), unit of information transmission which becomes subject of selection process (Wilkins, 1998), item of memory or a number of certain information that is stored neurally (Lynch, 1988). In memetics, we consider social system or culture as a system compounded by a great deal of units of cultural evolution or the smallest selection units that called meme (Wilkins, 1998).

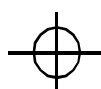
Artifacts, is one of cultural objects. It is a system constituted by cultural information units as the heredity units passes among generations in their evolution (read: innovation). We can determine the character of technological artifact by those information units expressing in particular way (Stankiewicz, 2000). In other words, artifact is a phenotypic meme or we call it phemetype emerged from its 'genotype', that is memetype. So we can say that the process innovation is somewhat a memetic evolution.

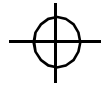
Innovation in memetic perspective can be regarded as a random mutation from meme codes of artifacts. Mutation or the change of meme from one code to different code in artifact will cause variation in artifact as the source of material for the selection process.

It is a difficult effort to identify memes that constitute an artifact. Since meme here is an abstract thing. However, as Situngkir cited in (2004), we can view meme from the evolutionary cultural object as the smallest unit of information we can identify and use to explain the evolution process. Here, memes as the smallest unit of information expresses certain characters or traits in artifact, be it the design, the technology, the function, and other traits. We can identify meme through observing the traits or the properties of an artifact and then determine the smallest unit of information that can be used to explain its evolutionary process. We took for example: design of cellphone. We can identify the meme of cellphone through certain character or trait in cellphone, i.e. design, function, technology, and other trait of cellphones, and then determine what the smallest unit of information expressing these traits or character which involve in its evolution.

3. Modeling Evolution of Technological Artifact through Memetic

In modeling artifact evolution in memetic perspective, we use a model developed in Situngkir (2004) and Situngkir, *et. al.* (2004). In the model, we describe artifact as a system constituted of





meme or unit of information as the smallest unit of replication. The memes will compound memplexes, where in memetic process expressed as phemotype i.e. the trait of design of artifacts. By denoting the set of memotype as M , phemotype as C , and function ρ as the function that correlates M with C , we formulate the relation as:

$$\rho: M \rightarrow C \quad (1)$$

Generally, memotype is a memplex constituted by a number of certain memes. Say it a memplex constituted of N memes, where each meme will have one alternative meme called allomeme (A), thus we can denote memplex as:

$$M = A_1 A_2 \dots A_N \quad (2)$$

with A assumed as set of all allomeme:

$$A = \bigcup_{i=1}^n A_i \quad (3)$$

As proposed by Heylighen (1993 and applied computationally by Situngkir (2004) and Situngkir *et al* (2004), allomeme in artifact can be stated as a "yes" or "no" propositions of certain character of artifact that made by the statement of "IF...THEN...". We then represent allomeme of each artifact as binary number (1,0) representing the presence of certain meme in the artifact.

Mutation process in a memplex of artifact will result a generation that has different set of memplex, which emerges as an artifact with different trait. In formal writing, this process is written as follow:

$$I: M \times C_M \rightarrow M^* \quad (4)$$

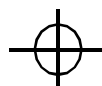
with I as mutation process in memplex that map its ancestor memplex M to the next generation memplex M^* with control parameter of mutation C_M

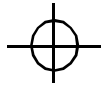
For more detail, in example of cellphone, we can use information of cellphone features as the basic information, which differ one type of cellphone to another.

The possible memes are:

1. IF it is infrared cellphone THEN there will be infrared system
2. IF it has SMS (short message service) function THEN there will be SMS technology
3. And so forth.

There will be a lot of memes that represent the characters of a specific artifact depends on available data. In the previous paper (Khanafiah & Situngkir, 2004), we have used cellphone Nokia for a case of study. We based the arrangement of its memplex on the feature information of each type. We picked 66 types of cellphone and in turn modeled the memplex for each type based on feature information of each variant. Thus, we identified 84 characteristic, be it design, or technological function that can possibly be the constitute memes of the memplex set of cellphone. The list of used allomemes and cellphone types can be seen in the referred paper.





4. Inference of Evolutionary Relationship of Technological Artifacts

Evolution is a gradual process, where simple species evolves becoming more complex species through accumulation of changes inherited over generations. A descendent will have several different properties from its ancestor because it is somehow changing in its evolution (Estabrook, 1987). Systematics is a branch in biology that studies genealogical relationship among organisms and also tries to describe the pattern of evolutionary events causing certain distribution and diversity in living things. Systematic analysis is conducted through constructing the history of evolution and the evolutionary relationship between descendents to their ancestors based on the similarities of characters as the basic of comparison (Lipscomb, 1998). This kind of analysis is well known as phylogenetic analysis, or sometimes called cladistics, which means a 'clade' or set of descendents from one common ancestor. Phylogenetic analysis is often represented as a branching system, a tree-like diagram known as phylogenetic tree (Brinkman, 2001).

Evolution of technological artifacts is of course different with the evolution in biological system. However, the same paradigm may become an alternative analysis to understand the innovation process in artifacts.

4.1. Phylogenetic Tree

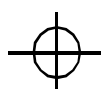
In living systems, evolutionary process involves genetic mutation and recombination process in a species so that it yields a new different species. Evolutionary history of an organism can be identified from the changes of its characters. The similar characteristics are the basis to analyzing the relationship of one species from other species.

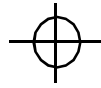
In this case, tree diagram is a logical way to show the evolutionary relationship among organisms (Schmidt, 2003). Phylogenetic is a model that represents the vicinity in ancestral relationship of organisms, the sequence molecules, or both (Brinkman, *et al*, 2001). Phylogenetic of organisms is the diagram that represents continuity of genealogy of an organism from over time, where a point of a branch showing a divergence while number of lines on certain times represents number of taxa at the interval time.

The arrangement of phylogenetic tree has several objectives, they are: to construct precisely genealogical relationship among organisms and to estimate the event where divergence occurs from one ancestor to their descendents (Li, *et al*, 1999).

In a form of graph, we formally define tree as an acyclic graph $T = (V, E)$, where V is set of vertices and E is set of edges set connecting two nodes of the graph. Phylogenetic as binary tree is a graph which has all nodes of degree one or three (Waterman, 1995). The degree of a node is defined as number of lines connected to the node.

In evolutionary tree (Figure 1), each node will represent a species or taxa. The lines or we call them braches depict the evolutionary development of the nodes. The nodes in the tree is divided into 3 types, they are root node, internal node, and external node. *Root node* $\in V$ is described as the node where the lines are in outward direction. External node or *leaf node* is the node representing species or taxa. This node has a degree of node = 1. Internal node is the node representing the common ancestor of each descendent. Internal node has a degree > 1 . In binary tree, internal node will have a degree = 3.





4.2. *Phylomemetic Tree*

As we have explained previously, inference or estimation of evolutionary history of an organism is one of the main goals in phylogenetic analysis. The similar paradigm in phylogenetic studies has become one of the most interesting alternatives to be proposed to analyzing artifact innovation as an evolving entity, in this case through memetics. The characteristics of a type of artifact determine the constituting memes. Innovation process as a random mutation from its memplex (eq. 4) yields a new kind of artifact that has several different “traits” with its ancestors. In short, we can use memes of artifact as the basics in understanding evolutionary relationship among artifacts.

Similarly with phylogenetic, we can depict the approximated evolution history and the evolutionary relationship among artifacts. We define the tree as a phylomemetic tree when we use the binary sequence of memes. Phylomemetic tree is a tree-like diagram describing evolutionary history and the relationship of the observing phenotypes or memplexes.

The phylomemetic tree for N types of artifacts with memplex M_i where $i = 1, \dots, N$, is an acyclic graph $T(N) = (V, E, M)$ consists of set of nodes V and set of edges E connecting one node with memplex M_i , $i = 1, \dots, N$, to the other. Each external node at phylomemetic tree represents artifact of which evolutionary relationship we tried to infer.

4.3. *Construction of Phylomemetic Tree*

As we mentioned previously, constructing phylomemetic tree is a way to infer the evolutionary relationship between artifacts. A phylomemetic tree is expected to describe evolutionary history of artifacts precisely enough based on the input data in form of sequence of memplexes for each artifact. That is why we need the right method of constructing phylomemetic tree.

There have been many methods used in constructing phylogenetic tree and we can use it as an alternative method to construct artifact phylomemetic tree. In this paper, we see several general

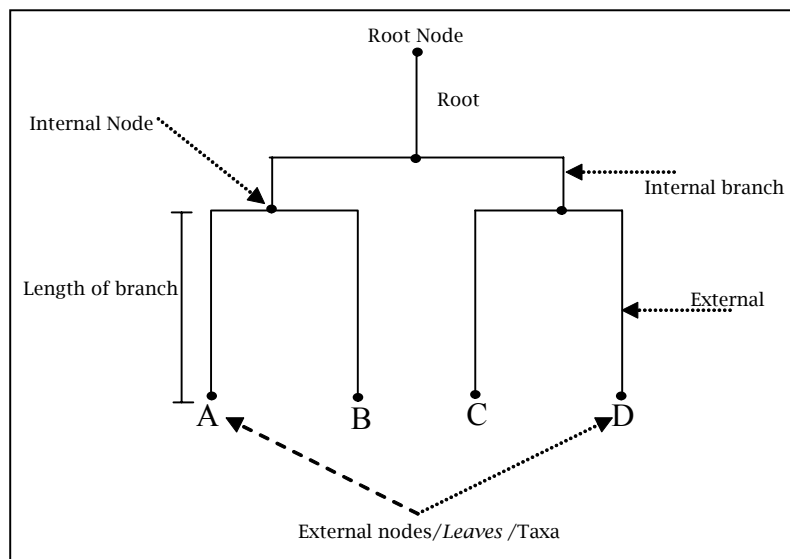
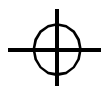
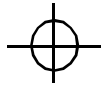


Figure 1 **Rooted phylogenetic tree**





methods in constructing phylogenetic tree to find out how the similar method used in constructing phylomemetic tree. Phylogenetic algorithm as widely known consists of two types, i.e.: (1) Algorithm construction using distance, including UPGMA and *Neighbor Joining* and (2) Algorithm construction based on character of: *parsimony maximum* and maximum likelihood (Brinkman, 2001).

In advance, the paper chooses 2 kinds of method in constructing phylogenetic, that is UPGMA and The Shortest Tree Method using Minimum Spanning Tree (MST) technique.

4.3.1. Unweighted Pair Group Method with Arithmetic Mean (UPGMA)

UPGMA or Unweighted Pair Group Method with Aritmatic Mean is a relatively simple and common method in constructing phylogenetic tree. It was developed for the first time to describe taxonomy phenogram i.e. a tree that describes phenotypic similarity among species, which then further developed as method for construction of phylogenetic tree under assumption that each species evolves independently with the same rate (Oppendoes, 1997).

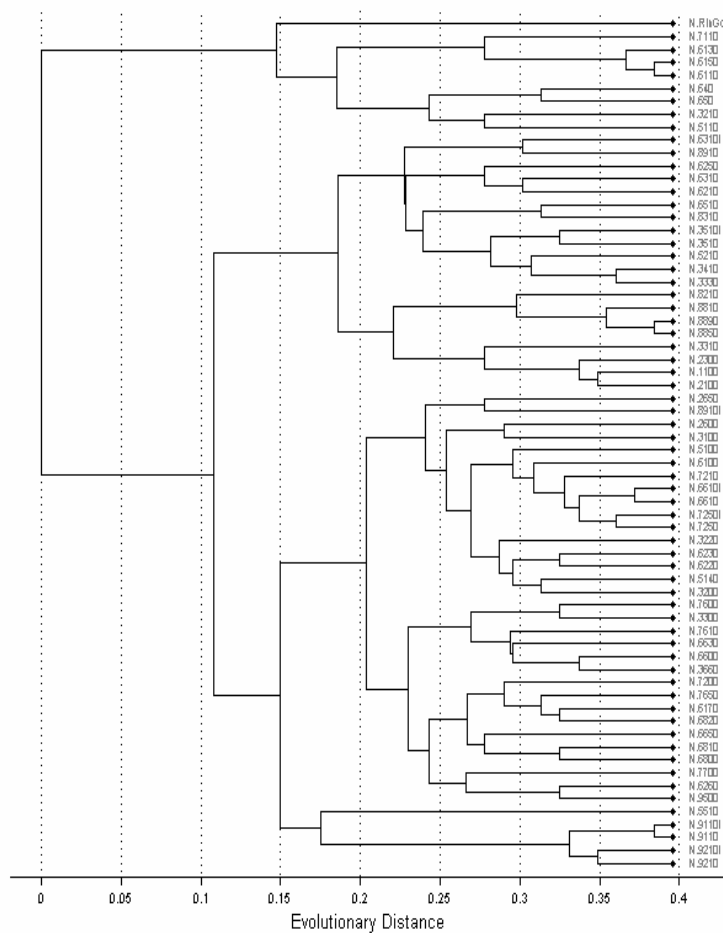
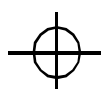
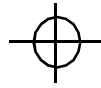


Figure 2 Example of the phylomemetic tree of Nokia cellphone designs obtained by using UPGMA





UPGMA is a tree-constructing algorithm using clustering analysis technique, which based on proximity among units to be represented as distance. When we choose genetic sequences, macromolecules or other sequences as the basic comparison of characters, then the distance between sequences will represent the cost mutation from one sequence to other sequences (Waterman, 1995. pp. 192).

UPGMA tries to build a tree where species correlate with the other with bigger character similarities, or has the minimum mutation distance. In UPGMA, we construct tree based on distances, where species with near distance is put in the same cluster. The distance between node of cluster c - that contains i and j with other unit k , is calculated using average system in arithmetic as follow:

$$d(c,k) = \frac{d(i,k) + d(j,k)}{2}; c = \{i, j\} \quad (5)$$

In short, the construction algorithm using UPGMA (Oppendoes, 1997) is as follows:

1. Choose the minimum distance between two taxa i and j , namely $d(i, j)$
2. Joint i and j as one new node $\{i, j\} = c$
3. Calculate distance from c to other node k using eq. 5
4. Erase column and line that contains i and j then substitute with $c = \{i, j\}$
5. Repeat from step 1 until last one row and column

UPGMA will yield the tree with topology or certain pattern of branching, which is expected to have the minimum distance between nodes. This is a quite fast method to approximate the topology of phylogenetic tree. However, it has a weakness since it does not include time evolution into account that much. Thus in general, the time when the species diverge into new species remain unknown.

The memplex of respective artifact can be used as a way to distinguish different characters or properties. To measure the distance between two memplexes, we use *divergence distance* between two sequences (Wagner, 1984), called *Manhattan distance* (Brooks, 1984). Manhattan distance is simply

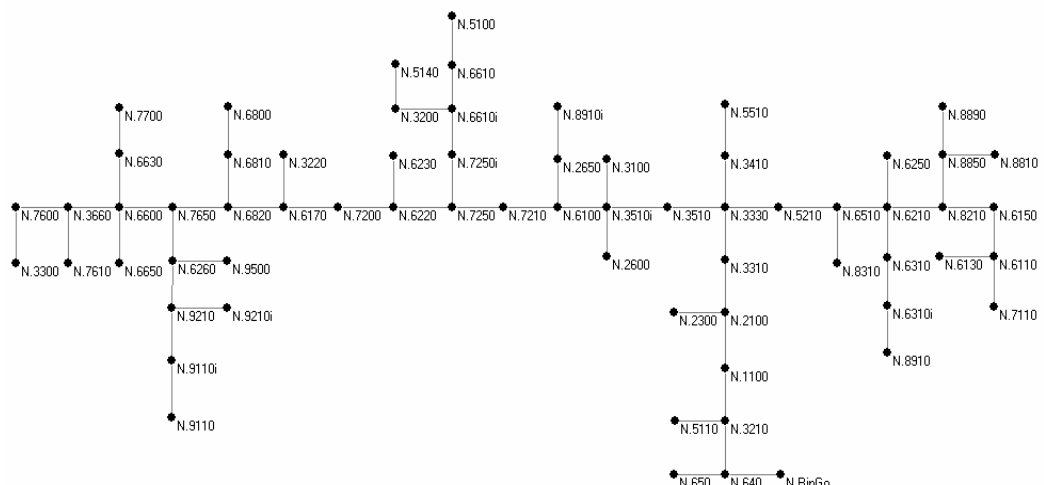
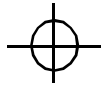


Figure 3

The Shortest Tree of Nokia cellphone yielded by the Minimum Spanning Tree of Kruskal Algorithm





defined as numbers of different characters between two sequences divided by total number of characters.

Regarding mutation value or cost of each meme representing esenting the change of state for each meme from 1 to 0 or vice versa, distance between artifact i and j with memplex $M_i = a_1 a_2 a_3 \dots a_k$ and $M_j = b_1 b_2 b_3 \dots b_n$ which is sequence with option (1,0) along k will value:

$$d(i, j) = \frac{\text{sum } a_i \neq b_i; i = 1, \dots, k.}{k} \tag{6}$$

The yielded distances become input to build our phylomemetic tree using UPGMA method as explained previously. Figure 2 shows an implementation of UPGMA to obtain the phylomemetic tree of Nokia cellphone designs.

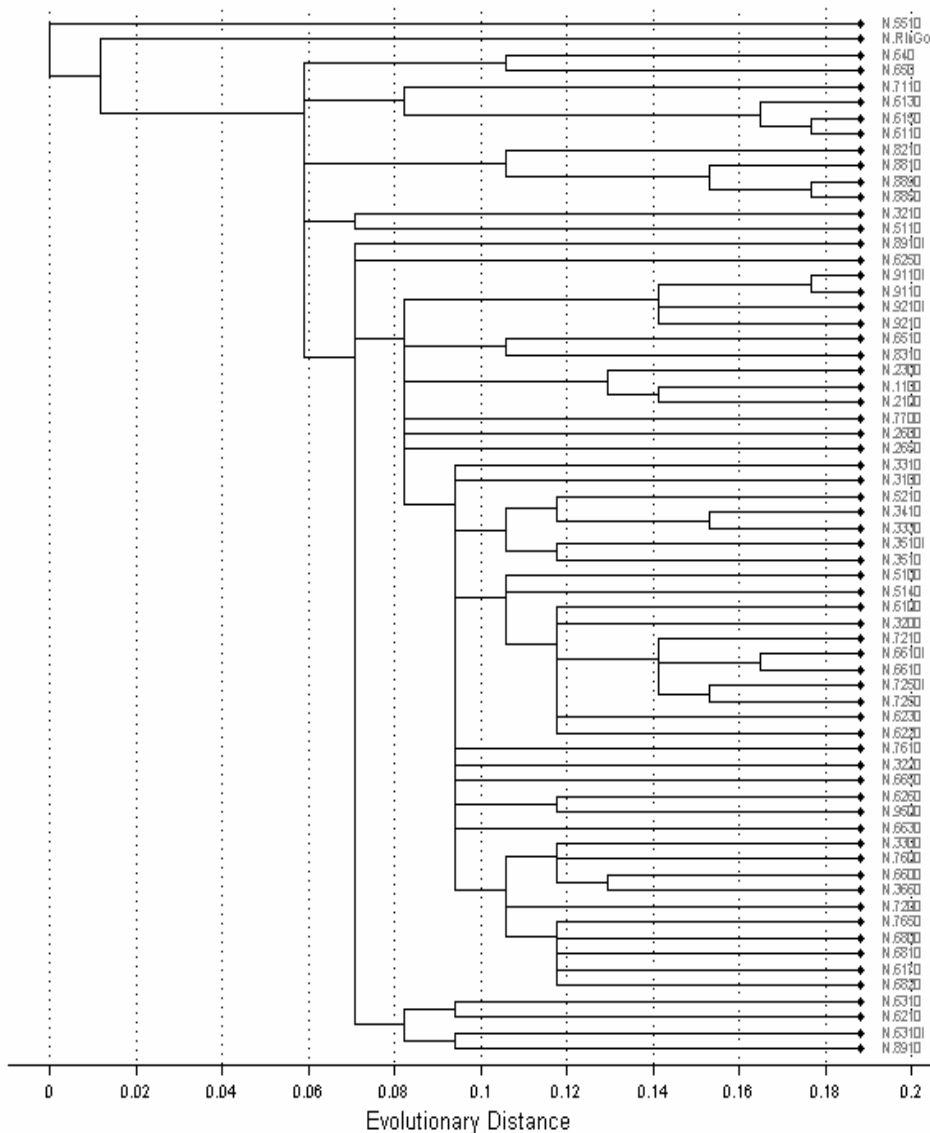
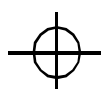
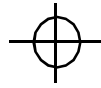


Figure 4 Example of the phylomemetic tree of Nokia cellphone designs obtained by using the Shortest Tree Algorithm





4.3.2. The Shortest Tree (Maximum Parsimony)

Discussions about phylogenetic tree are also frequently involving the notion of Hamming distance. Generally speaking, Hamming distance shows how much changes occurred upon two character sequences. The Hamming Distance between two sequences x and y with the same length is the number of different state for each same position for the two sequences. We can write it as $H(x,y) = |\{i : x_i \neq y_i\}|$. However, in the case of the paper, normalized Hamming distance is viewed as another word for *Manhattan* distance, having mutation value = 1 as we have previously explained. Thus, the Hamming distance between two memplexes is the number of different allomeme between the two sequences. In advance, we normalize it by dividing it with the number of memes constituting its memplex. The shortest distance of a tree is defined as minimum number of Hamming distance from the sequences constituting the tree.

For the sake of simplicity, we use *minimum spanning tree (MST)* technique by using Kruskal algorithm that approximately will give the minimum Hamming distance (Kruskal, 1956). MST is an algorithm to find the shortest path to connect one object with other objects in a system with certain number objects. The expected output is a tree that presumably gives the tree of summation of all distances is the smallest (minimum).

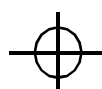
The minimum spanning tree is related to a subject in physics regarding a special subdominant ultrametric space with unique properties (See Mezard, *et. al.*, 2002). However, we have recognized a lot of use of ultrametric space of the minimally spanning tree in wide areas, e.g.: financial analysis (Situngkir & Surya, 2005). Interestingly, by applying the MST into the distances among artifacts, we could have geometric space depicting how the artifacts referring to each other. An example of this to the Nokia cellphone is depicted in figure 3. From this figure, we can see how an artifact (represented of the vertices) connected to others in the sense of minimally spanning tree that is recognized as the innovative reference among artifacts.

5. Discussions and Concluding Remarks

As discussed in detail in Khanafiah & Situngkir (2004), we could perform phylomemetic analysis with most innovated artifacts by using the two of methodologies. Interestingly, the first method is demonstrated to present visual estimation of how distant a series of artifact related to the other. Simply we can do this by looking at which type of the artifact branched from the same point. In the other hand, the tree generated from second method of The Shortest Tree gives an interesting picture of how a specific series of design referred to the other. In evolutionary speaking, we could interpret this as the relation of ancestor and descendant over time of innovation. The use of the phylomemetic analysis in cellphone designs is depicted in Figure 2 and 4, of UPGMA and The Shortest Tree respectively.

One important problem of constructing the phylogenetic analysis as also consequently discovered in phylomemetics is that of choosing certain allomemes used to build the tree. Apparently the way to decide the allomemes took into account for modeling the tree is notwithstanding subjective. Admittedly, this made us to note that we should give an extra careful attention and observation to choose any type of allomeme used in practical analysis.

Eventually we can remark several thoughts about innovation, that innovation is somewhat an





evolutionary process. We have shown the way to interpret the process of innovation by using memetics. In advance, we propose a methodology of phylomemetic analysis to visualize how the process possibly occurs. The two methods of UPGMA and the Shortest Tree show interesting features that can be valuable in the on going process of innovation: on the relative closeness between particular artifacts and how a design of an artifact refers to the other.

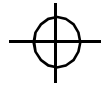
Acknowledgements

DK would like to give sincere thanks to Saras for literary assistance.

References

- Blackmore, S. (1998). 'Imitation and Definition of a Meme.' *Journal of Memetics - Evolutionary Models of Information Transmission*, 2.
http://jom-emit.cfpm.org/1998/vol2/blackmore_s.html
- Brinkman, F. & D. Leipe. (2001). 'Phylogenetic Analysis.' In A.D. Baxevanis & B.F.F. Ouellette., (eds.). *Bioinformatics: A Practical Guide to the Analysis of Gene and Protein*: 323-58. John Willey & Sons.
- Brooks, D. (1984). *Quantitative Parsimony*. In T. Duncan & T. Stuessy., (eds.). *Cladistics: Perspectives on the Reconstruction Evolutionary History*:119-32. Columbia University Press.
- Dawkins R. (1976). *The selfish gene*. Oxford University Press.
- Estabrook, G. (1984). 'Phylogenetic Trees and Character-State Trees.' In T. Duncan & T. Stuessy., (eds.). *Cladistics: Perspectives on the Reconstruction Evolutionary History*: 135-51. Columbia University Press.
- Felsenstein, J. (1984). 'The Statistical Approach to Inferring Evolutionary Trees and What It Tells Us About Parsimony and Compatibility.' In T. Duncan & T. Stuessy., (eds.). *Cladistics: Perspectives on the Reconstruction Evolutionary History* : 169-91. Columbia University Press.
- Frenken, K. (2001a). 'Modelling the organisation of innovative activity using the NK-model.' *Paper prepared for the Nelson-and-Winter Conference, Aalborg, 12-16 June 2001*.
<http://www.druid.dk/conferences/nw/paper1/frenken.pdf>
- Frenken, K. (2001b). *Understanding Product Innovation using Complex Systems Theory*. Academic Thesis. University of Amsterdam en cotuelle avec. Amsterdam.
<http://econ.geog.uu.nl/frenken/Dissertation%20Frenken.doc>
- Geser, H. (2004). *Toward a Sociological Theory of The Mobile Phone*. Publication online.
http://socio.ch/mobile/t_geser1.pdf
- Heylighen F. (1993). 'Selection Criteria for the Evolution of Knowledge.' *Proceeding 13th International Congress on Cybernetics, International Association of Cybernetics*: 524-8.
- Humphreys, L. (2003). *Can You Hear Me Now? A field study of mobile phone usage in public space*. University of Pennsylvania.
<http://www.asc.upenn.edu/usr/lhumphreys/PublicCellphones1.pdf>
- Joel, T., I. McCarty & D. Tranfield. (2002). 'A Structural and Evolutionary Approach to Change Management.' *Computational & Mathematical Organization Theory*, 8: 337-64.
- Kaplan, S. & M. Tripsas. (2003). 'Thinking about technology: understanding the role of cognition and





- technical change.' *Working Paper Series December 2003. Industrial Performance Center. Massachusetts Institute of Technology.*
<http://web.mit.edu/ipc/www/pubs/articles/03-008.pdf>
- Khanafiah, D. & H. Situngkir. (2004). 'Innovation as Evolution - Phylomemetic of Cellphone Designs.' *Working Paper WPV2004, Bandung Fe Institute.*
- Kauffman, S. (1995). *At Home in The Universe: The Search for Laws of Self-Organization and Complexity.* New York: Oxford University Press.
- Kruskal, J. (1956). 'On the shortest spanning subtree of a graph and the travelling salesman problem.' *Proceeding of American Mathematical Society*, 7: 48-50.
- Lipscomb, D. (1998). 'Basics of Cladistic Analysis.' *Student guide paper. George Washington University.*
<http://www.gwu.edu/~clade/faculty/lipscomb/Cladistics.pdf>
- Li, S., D. Pearl & H. Doss. (1999). 'Phylogenetic Tree Construction using Markov Chain Monte Carlo.' *Fred Hutchinson Cancer Research Center Washington.*
<http://www.stat.ohio-state.edu/~doss/Research/mc-trees.pdf>
- Lynch, A. (1998). 'Units, Events and Dynamics in Memetic Evolution'. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2.
http://jom-emit.cfpm.org/1998/vol2/lynch_a.html
- Mezard, M., G. Parisi & M. Virasoro. (eds.). (2002). *Spin Glass Theory and Beyond.* World Scientific.
- Mokyr, J. (1997). 'Innovation and Selection in Evolutionary Models of Technology: Some Definitional Issues.' *Paper prepared for the Conference on Evolutionary Models in Economics, Jan. 9-12, 1997.*
- Opperdoes, F. (1997). 'Construction of a distance tree using clustering with the Unweighted Pair Group Method with Arithmetic Mean (UPGMA).' Online publication.
<http://www.icp.ucl.ac.be/~opperd/private/upgma.html>
- Sartika, T. (2004). 'Tracing Cultural Evolution Through Memetics.' *Working Paper WPF2004, Bandung Fe Institute.*
- Schmidt, H. (2003). 'Phylogenetic Trees from Large Datasets.' *Inaugural-Dissertation, Dusseldorf University.*
<http://www.bi.uniduesseldorf.de/~hschmidt/publ/schmidt2003.phdthesis.pdf>
- Situngkir, H. (2004). 'On Selfish Meme: Culture as Complex Adaptive System.' *Journal of Social Complexity*, 2(1): 20-32.
- Situngkir, H. & Y. Surya. (2005). 'On Stock Market Dynamics through Ultrametricity of Minimum Spanning Tree.' *Working Paper WPH2005, Bandung Fe Institute.*
- Situngkir, H., T. Sartika & D. Khanafiah. (2004). 'Stabil Evolusioner Partai Politik di Indonesia - Studi Kasus: Menjelang Pemilu 2004.' *Working Paper Series WPH2004, Bandung Fe Institute.*
- Stankiewicz, R. (2000). 'The Concept of "Design Space".' *Paper of Research Policy Institute, University of Lund, Sweden.*
- Waterman, M. (1995). *Introduction to Computational Biology: Maps, Sequence and Genomes - Interdisciplinary Statistics: 345-67.* London: Chapman & Hall.
- Wagner, W. (1984). 'Application of the Concept of Groundplan-Divergence.' In T. Duncan & T. Stuessy., (eds.). *Cladistics: Perspectives on the Reconstruction Evolutionary History.* 95-118. Columbia University Press.
- Wilkins, J. (1998). 'What's in a Meme? Reflections from the perspective of the history and philosophy of evolutionary biology.' *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html

