ESD 342 class 14 decomposition

C. L. Magee

March 30, 2010

Class 14 Lecture outline

- ☐ **Decomposition**
  - ■ Link to modularity
  - ■ Practical and theoretical importance
  - ■ Taxonomy and examples
  - ■ Network-based Approaches to Quantitative Decomposition
    - ☐ Structural or cohesive decomposition
    - ☐ Functional decomposition
      - ■ Roles, positions and hierarchy
      - ■ Motifs and course graining
- ☐ **Overview of modeling**

MIT ESD
Massachusetts Institute of Technology
Engineering Systems Division

MIT

Decomposition, modularity and taxonomy

- ☐ What if anything relates these concepts, what is different?

- ☐ Breaking the problem into smaller parts and "organizing" in an intelligent manner?

- ☐ Some narrower definitions

  - ■ Modularity can be restricted to systems where "plug and play" is possible –this combines in a restrictive way two types of modularity 1 and 2 after Whitney

Better Definition(s) of Modularity- Whitney

- ☐ Modularity 1:
  - ■ The system **can be decomposed into subunits** (to arbitrary depth)
  - ■ These subunits can be dealt with separately (to some degree)
    - ☐ **In different domains, such as design, manufacturing, use, recycling**
- ☐ Modularity 2:
  - ■ The functions of the system can be associated with clusters of distinct elements
    - ☐ **in the limit one function:one module**
  - ■ These elements operate somewhat independently
  - ■ They do not have to be physically contiguous
- ☐ Common to both definitions
  - ■ Independence of some kind
  - ■ Identifiable interfaces (perhaps standardized)
  - ■ More interactions inside a module, fewer interactions between modules

Decomposition, modularity and taxonomy

- ☐ What if anything relates these concepts
- ☐ Breaking the problem into smaller parts and "organizing" in an intelligent manner?
- ☐ Some narrower definitions (however it is not possible to dictate wide use of these now)
  - → ■ Modularity can be restricted to systems where "plug and play" is possible –this combines in a restrictive way modularity 1 and 2
  - ■ "Plug and Play" Modularity also aligns decomposability in different domains- design, manufacturing, re-use, recycling, operation
  - ■ Plug and Play modularity is quite rare and is the opposite of integrality

Integrality and Modularity (2) in Engineered Systems

- ☐ **Modular (2) systems are, ideally, those in which**
  - ■ Functions and behaviors can be associated simply and directly with modules more or less one-to-one
  - ■ Only predefined interactions occur between modules
  - ■ Interactions occur at, and only at, predefined interfaces
  - ■ Modules don't need to know what is on the other side of the interface

- ☐ **Integral systems differ as follows:**
  - ■ Functions are shared among modules
  - ■ It matters what is on the other side of the interface
  - ■ Interactions that were not defined can occur, and they can occur at undeclared interfaces
  - ■ Behaviors can arise that are not easily traceable to modules one-to-one
  - ■ In many cases you can't stop this from happening
  - ■ To the extent that this occurs, all systems are more or less integral

- ☐ **"Integral" and "modular" represent extremes and all real systems lie in between**

What Makes Something "Inherently Integral?" (Whitney)

- ☐ It has multiple performance attributes (a measure of complexity?)
- ☐ Attribute delivery is distributed within the product/system, and shared by many parts
- ☐ The attributes are coupled and may conflict
  - ■ car door leaks helped by tight seals
  - ■ car door closing effort hurt by tight seals
- ☐ Inter-module couplings are very strong
  - ■ load paths in aircraft structure
  - ■ data exchanges in time-critical computing tasks
- ☐ As a result, the product/system may appear modular but it is not

Decomposition, modularity and taxonomy

- ☐ What if anything relates these concepts?
- ☐ Breaking the problem into smaller parts and "organizing" in an intelligent manner?
- ☐ Some narrower definitions (however it is not possible to dictate wide use of these now)
  - ■ Modularity can be restricted to systems where "plug and play" is possible
  - ■ Taxonomy can be restricted to the further breakdown (*or hierarchical aggregation*) of concepts
- ☐ In this lecture, I will use these restricted definitions and think of both as contained within the broader concept of *decomposition- attempting to reduce the complexity of an artifact, process, algorithm, enterprise, problem or concept by defining sub-groups (reductionism?)*.

Importance of decomposition

- ☐ From a practice perspective
  - ■ Decomposition is essential- problems are too big to not be broken into parts to think about them, to manage progress and to have small enough groups working to make progress at a decent rate (engineering, management and science)
  - ■ The interfaces are where the most difficult issues arise- one must pay particular attention to problems or attributes which do not really fit well into the decomposed activities or elements
  - ■ To minimize the number of difficult issues, one wants to decompose as effectively as possible- that is to end up with interfaces that are as simple as can be designed ("plug and play" is ideal but..).
- ☐ "Complications" and "opportunities"
  - ■ There are fundamental limits to decomposition (just discussed)
  - ■ There are **many ways to decompose** (or aggregate) both **strategically** and tactically
  - ■ One would also like to know how "integral" the system is even after decomposition. Most possible decompositions contain imperfections-can we **quantitatively assess** the lack of perfection?
  - ■ Can we objectively pick the best decomposition?

Decomposition and Taxonomy

☐ Taxonomy has "rules" for judging perfection of the decomposition. These rules are (almost) never found to be perfectly followed by our attempts to categorize but they might still guide one to more effective decompositions.

☐ The goals of a taxonomy

Collectively Exhaustive and *Mutually Exclusive* (*DIFFICULT*)

Internally Homogeneous (ALSO DIFFICULT)

Stability

Understandable Representation and Naming

Examples follow: first general and then from ESD342

Journal of Economic Literature Classification System (19 categories)

- General Economics and Teaching
- Schools of Economic Thought and Methodology
- Mathematical and Quantitative Methods
- Microeconomics
- Macroeconomics and Monetary Economics
- International Economics
- Financial Economics
- Public Economics
- Health, Education and Welfare
- Labor and Demographic Economics
- Law and Economics

- Industrial Organization
- Business Administration and Business Economics; Marketing; Accounting
- Economic History
- Economic Development, Technological Change, and Growth
- Economic Systems
- Agricultural and Natural Resource Economics: Environmental and Ecological Economics
- Urban, Rural and Regional Economics
- Other Special Topics

Journal of Economic Literature Classification System (19 categories)

- General Economics and Teaching
- Schools of Economic Thought and Methodology
- Mathematical and Quantitative Methods
- Microeconomics
- Macroeconomics and Monetary Economics
- International Economics
- Financial Economics
- Public Economics
- Health, Education and Welfare
- Labor and Demographic Economics
- Law and Economics

- **Industrial Organization**
- Business Administration and Business Economics; Marketing; Accounting
- Economic History
- Economic Development, Technological Change, and Growth
- Economic Systems
- Agricultural and Natural Resource Economics: Environmental and Ecological Economics
- Urban, Rural and Regional Economics
- Other Special Topics

Journal of Economic Literature Classification System: Typical Final Substructure in which books and papers are the next level of detail

- □ L. Industrial Organization (one of 19 highest level categories)
  - ■ L9 – Industry Studies: Transportation and Utilities
    - □ L90- General
    - □ L91- Transportation: General
    - □ L92- Railroads and Other Surface Transportation: Autos, Buses, Trucks and Water Carriers; Ports
    - □ L93 Air Transportation
    - □ L94- Electric Utilities
    - □ L95- Gas Utilities; Pipelines; Water Utilities
    - □ L96- Telecommunications
    - □ L97-Utilities: General
    - □ L98- Government Policy
    - □ L99- Other

Assessing a decomposition- JEL example

- ☐ The JEL categories do not come very close to achieving the Ideal rules for taxonomy. How would one assess whether it should be improved?
- ☐ Time and confusion about users reaching *correct* bins (relative to what user and compared to what?)
- ☐ Relative size of material collected in the individual bins? (20 Qs)
- ☐ Internal homogeneity at lowest level?
- ☐ Appropriate adaptation to new fields of study?

Decomposition and Taxonomy II

- [ ] Taxonomy has "rules" for judging perfection of the decomposition. These rules are (almost) never found to be perfectly followed by our attempts to categorize but they might still guide one to more effective decompositions.

- [ ] The goals of a taxonomy

  Collectively Exhaustive and *Mutually Exclusive* (*DIFFICULT*)

  Internally Homogeneous (ALSO DIFFICULT)

  Stability

  Understandable Representation and Naming

  → Plus: finite variation in group size and use of hierarchy possible to arrive at "Internal Homogeneity"

  Examples follow: first general and then from ESD342

MIT ESD
Massachusetts Institute of Technology
Engineering Systems Division

Structure in Biology

☐ **Sub-division of the field of study (e.g. publications)**

■ No parallel found to *JEL* Classification System

■ Textbook topics are at best approximately homologous with the activities in the field. Areas such as bioinformatics, mathematical modeling and observational techniques receive relatively more listings than for economics so would probably appear on an equivalent to the *JEL* Classification. Also, it is clear that systematics is an old and still very active field not much covered in modern undergraduate textbooks. Paleontology is not part of biology but is essential to Biology Systematics.

☐➤ **Sub-division (and aggregation) of the objects studied**

■ Extensive, historically dominant in field and still very active

Biological Classification History

- Aristotle identified hundreds of differing kinds of animals and plants based upon *morphology and function*.

- John Ray: 1628-1705, [Cambridge UK] restarted biological classification based on *morphology and structural similarity*.

- Linnaeus 1707-1778, [Swedish] is the most recognized biological classifier. He introduced the idea of a unified hierarchical tree (7 levels) and the "binomial" nomenclature for species and modified versions of both of these still stand. However, the characteristics he used for classification (*sexual reproduction modes and organs)* are now not as important as the ones used by Ray.

- Aggasiz, Paley, Sedgwick, Buckland (19th century): paleontology, embryology, ecology, and biogeography all became important in classification through their work.

- Robert Whitaker in 1969 proposed 5 kingdoms whereas Linnaeus only had 2 (plants and animals).

Biological Classification

- 5 Kingdoms: Plants, Animals, Fungi, Protists, bacteria
- The 7 layer hierarchy continues:
  - Phylum (for animals) and Divisions (for plants and Fungi)
  - Class
  - Order
  - Family
  - Genus
  - Specie (named by genus + Latinized specific)
- The bottom 2 layers are fairly well specified but the middle is a real muddle and extremely hard to make sensible in all 5 kingdoms. Moreover, the top has added a layer in the past few decades (single vs multi-cell).
- Remains a very active area of biological science with constant modification and heavy use of genomes etc. with continued conflicts about proper categorization based upon characteristics of interest

Elemental Classification

☐ By mid-nineteenth century, 60 elements were known (1/5 discovered by Humphrey Davy between 1780 & 1825)

☐ Dalton (1808) and Avogadro (1811 but "lost" for 50 years) did key work defining atoms and their size but the array of elements with different characteristics encouraged many to attempt a rational classification system:

- ■ Elemental forms - gases, liquids, solids;
- ■ Elemental atomic weight (and number)
- ■ Elemental properties like hardness, conductivity, reflectivity
- ■ Elemental reactivity with various other elements to form stable compounds with various properties

☐ Mendeleyev in 1869 combined the two most popular classification schemes (atomic weight and common properties) into a single *table* (atomic number with a period of 7) and "Without a doubt, the Periodic Table of the Chemical Elements is the most elegant organizational chart ever achieved" and greatly clarified and stabilized Chemistry as a science and made atomic physics and other fields possible.

Systems *Context* Typology

- ☐ Overtly designed
  - ■ Can be an architect
  - ■ A design strategy is possible to imagine
  - ■ Products, product families
  - ■ Cars, airplanes
  - ■ Bell System
  - ■ Organizations
  - ■ Centrally-planned economies

- ☐ Partially designed
  - ■ Architect not common
  - ■ Protocols and standards are crucial
  - ■ Design strategy may not be practical
  - ■ May be designed in small increments
  - ■ Usually grow with less direction from a **common** strategy over longer times
  - ■ Regional electric grids
  - ■ City streets
  - ■ Federal highway system
  - ■ *MIT?*

- • Non-designed systems
  - – No architect
  - – Respond to context
  - – Change, develop
  - – Differentiate or speciate
  - – Interact hierarchically, synergistically, exploitatively
  - – Cells, organisms, food webs, ecological systems
  - – *Friendship groupings?*
  - – *Co-author networks?*

Structural Typology

- ☐ Totally regular
  - ■ Grids/crystals
  - ■ Pure Trees
  - ■ Layered trees
  - ■ Star graphs
- ☐ Deterministic methods used

- Real things
  - The ones we are interested in
- New methods or adaptations of existing methods needed

- No internal structure
  - Perfect gases
  - Crowds of people
  - Classical economics with invisible hand
- Stochastic methods used

- Less regular
  - "Hub and spokes"
  - "Small Worlds"
  - Communities
  - Clusters
  - Motifs

Operand / Operation	Matter (M)	Energy (E)	Information (I)
Transform	blast furnace	***Engines, Electric motors***	**Analytic engine, Calculator**
Transport	Truck	***Electrical Grid***	**Cables, Radio, Telephone and Internet**
Store	Warehouse	***Batteries, flywheels, Capacitors***	**Magnetic tape and disk, Book**
Exchange	eBay Trading System	Energy Markets	World wide web, Wikipedia
Control	Health Care System	Atomic Energy Commission	Internet engineering task force

Importance of decomposition II

- ☐ From a practice perspective
  - ■ Decomposition is essential- problems are too big to not be broken into parts to think about them, to manage progress and to have small enough groups working to make progress at a decent rate (engineering and science)
  - ■ The interfaces are where the most difficult issues arise- one must pay particular attention to problems or attributes which do not really fit well into the decomposed activities or elements
  - ■ To minimize the number of difficult issues, one wants to decompose as effectively as possible- that is to end up with interfaces that are as simple as can be designed
- ☐ Decomposition *rarely* (once?) *meets the ideal "rules of taxonomy" but is nonetheless useful* (if carefully used)
- ☐ Tools that can help
  - ■ Application of Hierarchy (recursive decomposition as in 20 questions)
  - ■ Bottom-up as well as top-down examination
  - ■ Application of Matrices (Elemental) often useful
  - ■ Characteristics for decomposition must be analyzed

Class 14 Lecture outline

- ☐ Decomposition
  - ■ Practical and theoretical importance
  - ■ Link to modularity
  - ■ Taxonomy and examples
  - ■ Network-based Approaches to Quantitative Decomposition
    - ☐ Structural or cohesive decomposition
    - ☐ Functional decomposition
      - ■ Roles, positions and hierarchy
   - ■ Motifs and course graining
- ☐ Overview of modeling

Steps toward quantitative decomposition based upon network models

- Systems to be decomposed are represented as networks among elements that have relationships indicated by links
- "**Strategic" Question**: What characteristics do we use to decide upon decomposition?
- We first consider only simple networks with one kind of node and one kind of link but even in this "simple" case, we will see several strategic ways (at least three) to logically decompose the system with different meanings and different answers to the **tactical questions**
- How many subgroups (and what members)?
- How "perfect" is the proposed decomposition?

Two strategically different approaches for decomposition of a network

□ First *quantitatively* pursued in Mo-Han Hsieh's thesis with application to decomposing the citation network of the Internet standards into meaningful subgroups but the basic ideas were developed by social network researchers 35 years ago

  ■ Cohesion to others in subgroup

  ■ Role similar to others in subgroup (hierarchy)
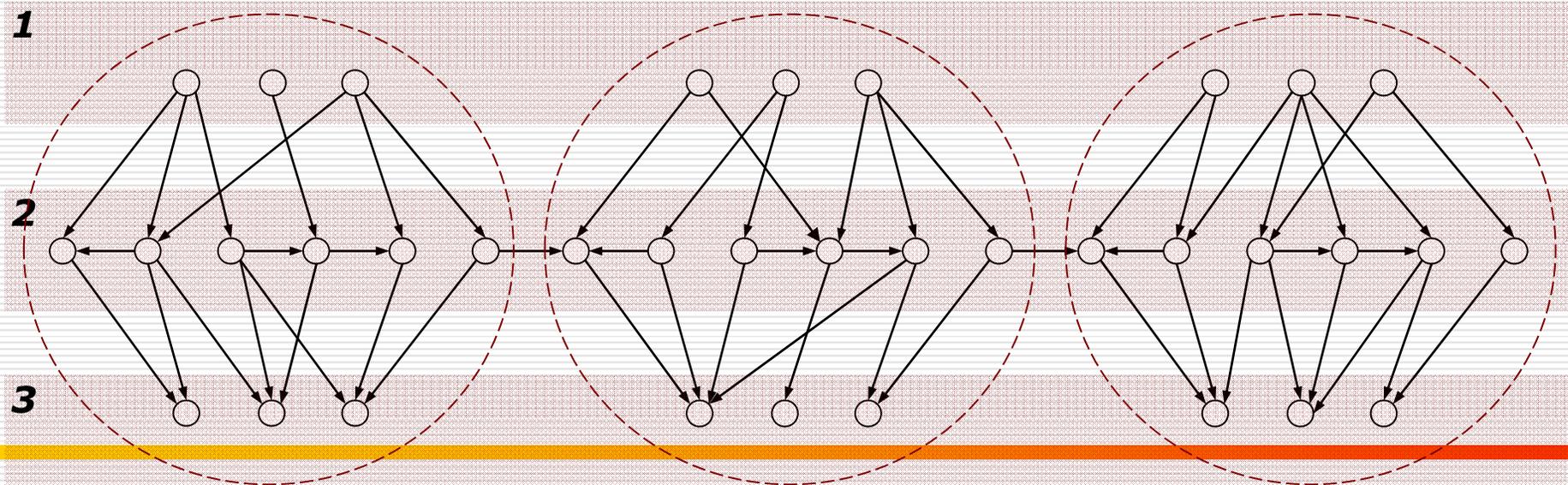
The tactical "answers: Algorithms for decomposition and decomposability metrics (tactics)

- ☐ *Cohesion*: Newman-Girvan algorithm and Newman modularity (and newly derived normalized decomposability metric)

- ☐ Role (position or hierarchy): Hsieh-Magee algorithms and decomposability metric for *structural* and *regular equivalence*

- ☐ All three concepts have been defined in the Social Network Literature

- ☐ Cohesive sub-groups are formed among nodes (agents) who have links among each other more often than with those in other sub-groups

Cohesive decomposition:
The Newman-Girvan algorithm

- ## The algorithm
  1. Calculate the betweenness of all edges in the network.
  2. Remove the edge with the highest betweenness.
  3. Recalculate the betweenness of all remaining edges.
  4. Repeat from step 2 until no edges remain (Max Q is best).

- ## The community structure (i.e. dendrogram)

- ## Modularity: Q *(To determine the best number of communities)*
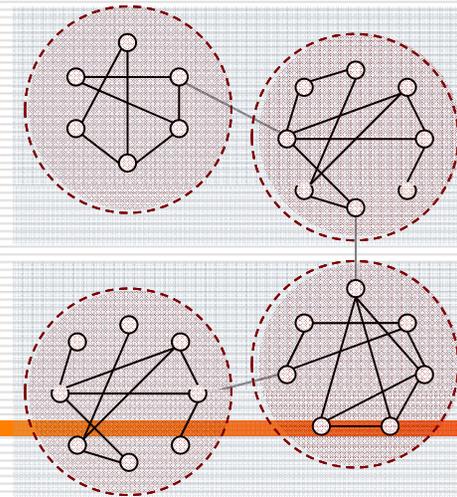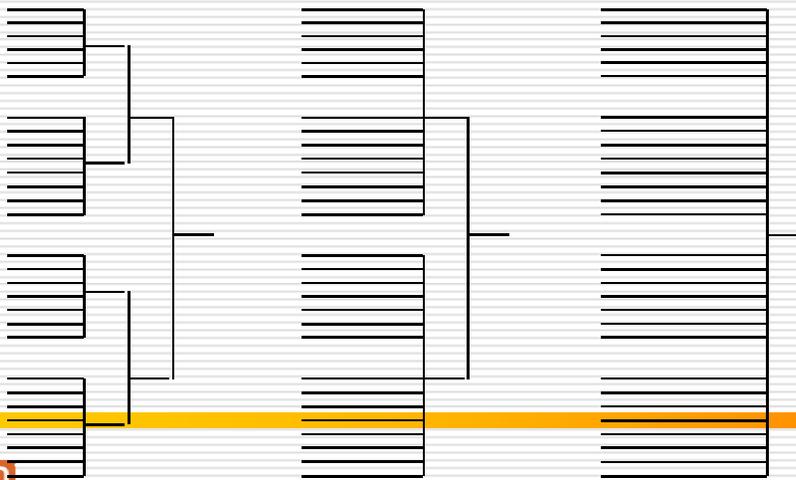
Newman Modularity Metric

- ☐ Basic idea
  - ■ the sum of the fraction of intra-group edges minus the value that it would take if edges were placed at random.

$$Q = \sum_i \left(e_{ii} - a_i^2 \right)$$

- ☐ e_{ij} - the fraction of edges in the network that connect vertices in group i to those in group j,
- ☐ a_i - is the fraction of all edges that go out from vertices in group i or come in to vertices in group i
  - ■ (i.e. $a_i = \sum_j e_{ij}$ or $a_i = \sum_i e_{ij}$)   (Newman 2004).
- ☐ This metric is used by Newman and Girvan as a "stopping rule"- the correct number (and members) of subgroups maximizes Q (answering the first tactical question for this type of decomposition)

Cohesive Decomposability Metric

- ☐ Q cannot be used to compare how effective a decomposition is between different networks
  - ■ To compare networks of different sizes, different numbers of sub-groups and different link densities, one needs a properly normalized metric
- ☐ Normalized Cohesive Decomposability Metric: Q_n
  - ■ Let p be the number of sub-groups
  - ■ Let n be the total number of edges of the network
  - ■ For a connected network, the largest possible fraction of intra-group edges: $f=1-(p-1)/n$
  - ■ We normalize the Newman modularity measure by f minus the value that it would take if edges were placed at random.

$$Q_n = \sum_i \left(e_{ii} - a_i^2 \right) \Big/ \left(f - \sum_i a_i^2 \right)$$

Cohesive Decomposability Metric: Example

$$e = \begin{bmatrix} 0.31 & 0.00 & 0.00 \\ 0.03 & 0.31 & 0.00 \\ 0.00 & 0.03 & 0.31 \end{bmatrix} \quad a = \begin{bmatrix} 0.31 \\ 0.34 \\ 0.34 \end{bmatrix}$$
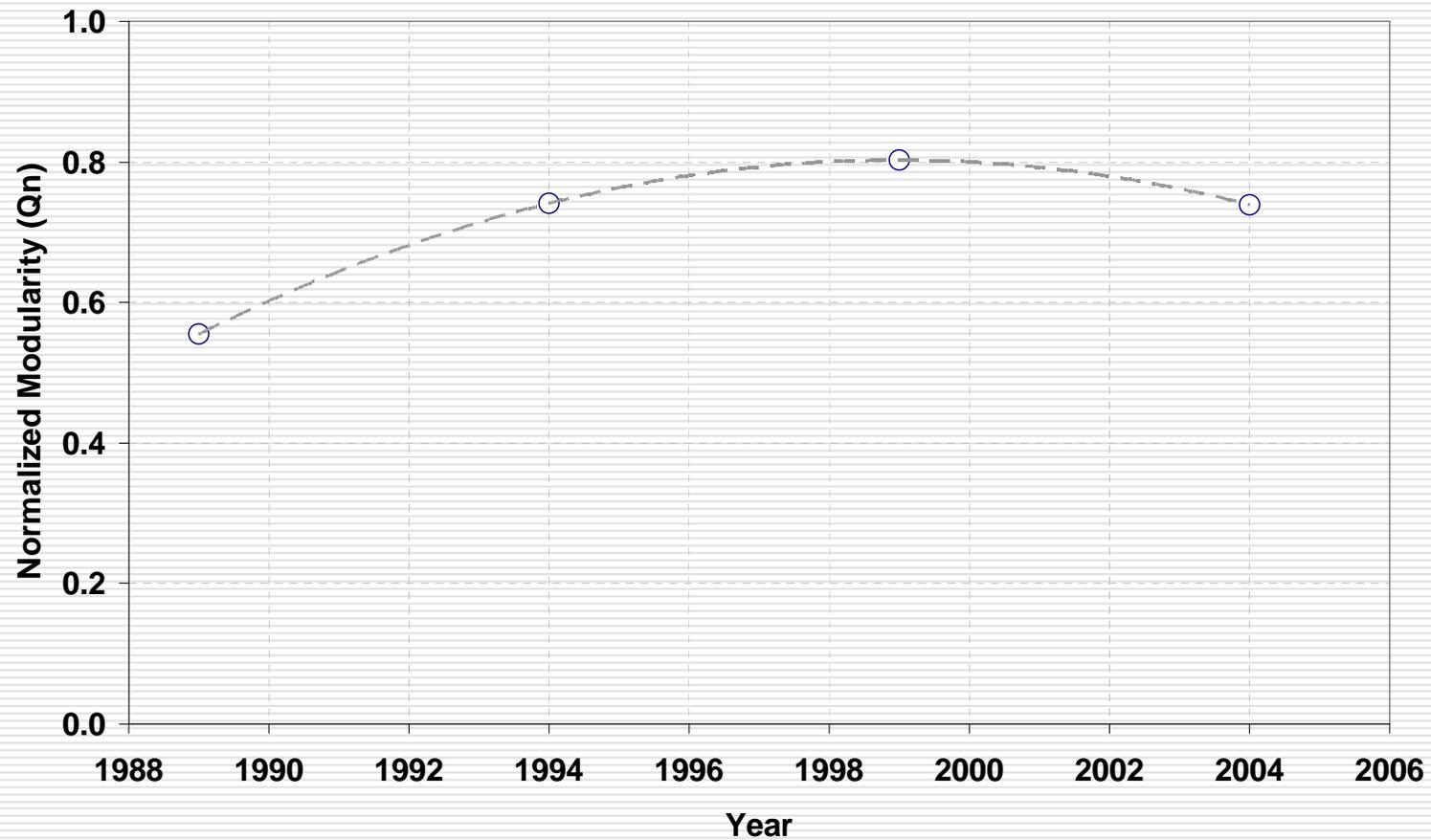
$$f = 1 - (p-1)/n$$
$$= 1 - (3-1)/35 = 0.93$$

$$Q = \sum_i \left(e_{ii} - a_i^2 \right) = (0.31 - 0.31^2) + (0.31 - 0.34^2) + (0.31 - 0.34^2) = 0.60$$

$$Q_n = Q/(f - \sum_i a_i^2) = 0.60 / \left[0.93 - (0.31^2 + 0.34^2 + 0.34^2) \right] = 1.00$$
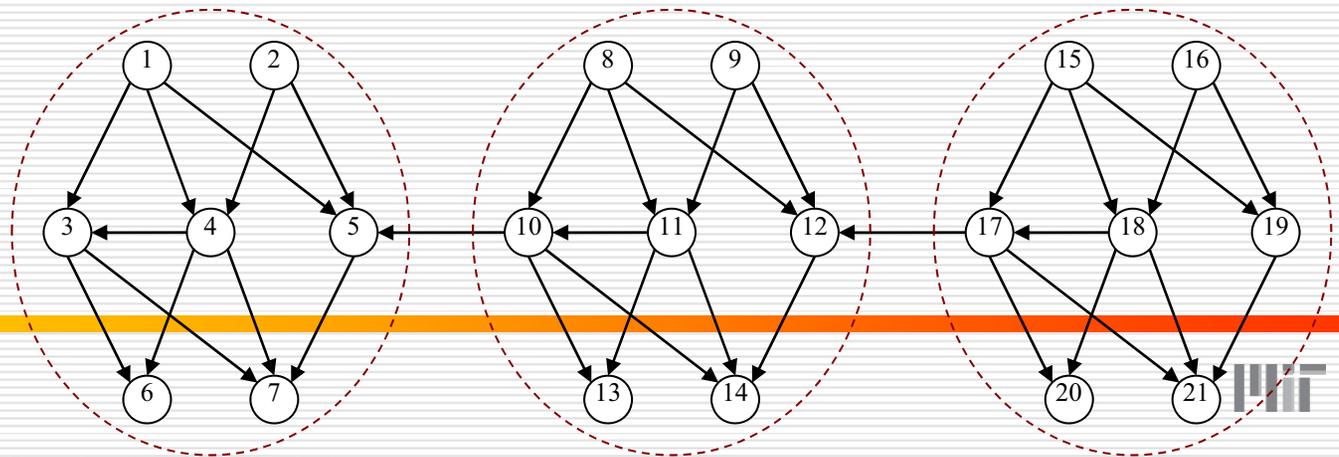
Cohesive Decomposability Metric Example – Internet Standards

Hsieh-Magee Approach to Decomposition for Structural Equivalence

☐ **Structural Equivalence**

  ■ if two nodes link to and are linked by exactly the same set of other nodes, they are structurally equivalent to each other.

  ■ E.g. Node 1 and 2 are structurally equivalent. Node 3, 4 and 5 are structurally equivalent
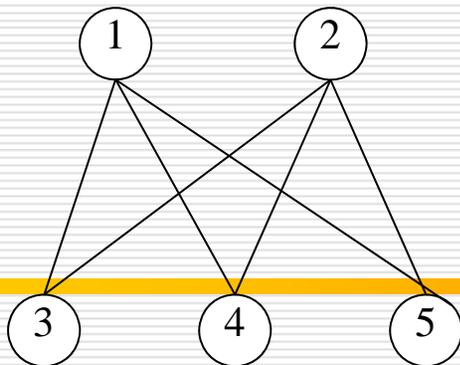
$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

# Hsieh-Magee Approach to decomposition for Regular Equivalence

□ **Regular Equivalence**
  - ■ Vertices that are regular equivalent are not generally connected to the same vertices, but the vertices they connect to are in the same class as each other.
  - ■ A regular block contains at least one arc in each row and in each column

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
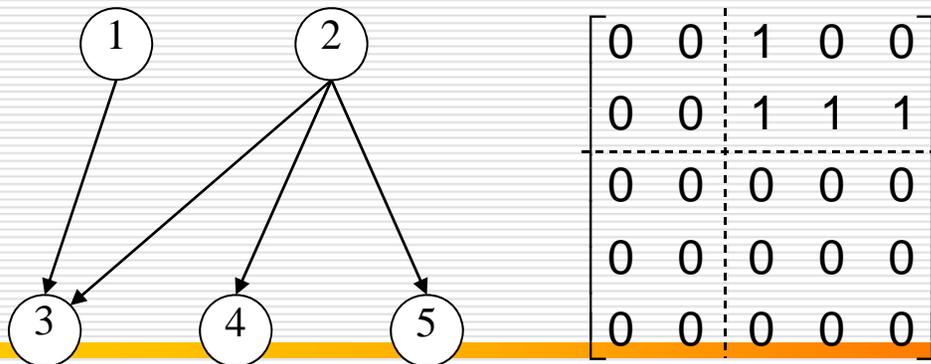
Two *strategically* different approaches for decomposition of a network II

- ☐ First *quantitatively* pursued in Mo-Han Hsieh's thesis with application to decomposing the citation network of the Internet standards into meaningful subgroups but the basic ideas were developed by social network researchers 35 years ago
  - ■ Cohesion to others in subgroup
  - ■ Role similar to others in subgroup (hierarchy)
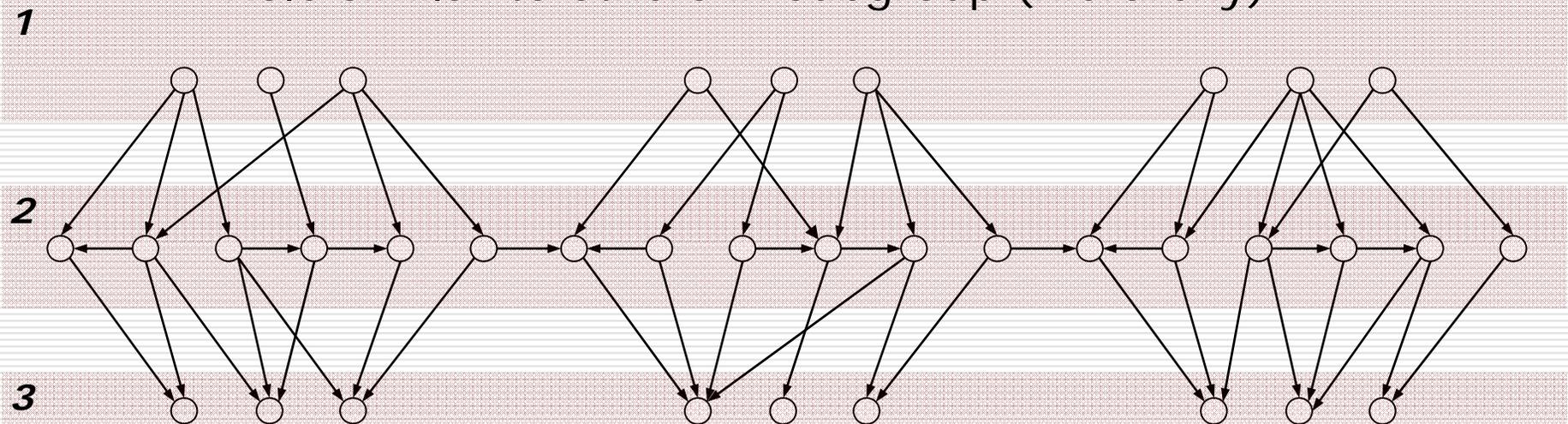
Decomposition by Role:
The algorithm

- The *algorithms* (and the decomposability metrics) for structural and regular equivalence are very similar:

- Transform *n* by *n* adjacency matrix into a *n* by *n* Similarity matrix by use of the definitions of structural (and regular) equivalence

- View *n* by *n* Similarity matrix as *n* nodes in *n* dimensional space

- Apply K means algorithm to find *k* sub-groups of nodes that best match (are most similar to) each other

- Use comparison to **random network changes** to arrive at best number and members of sub-groups (answers first tactical question for this strategic approach to decomposition)

MIT ESD
Massachusetts Institute of Technology
Engineering Systems Division

© 2007 Chris Magee, Engineering Systems Division, Massachusetts Institute of Technology

SE Decomposition applied to the inter-organizational Search and Rescue (SAR) network created after a disaster in Kansas (Drabek 1981)
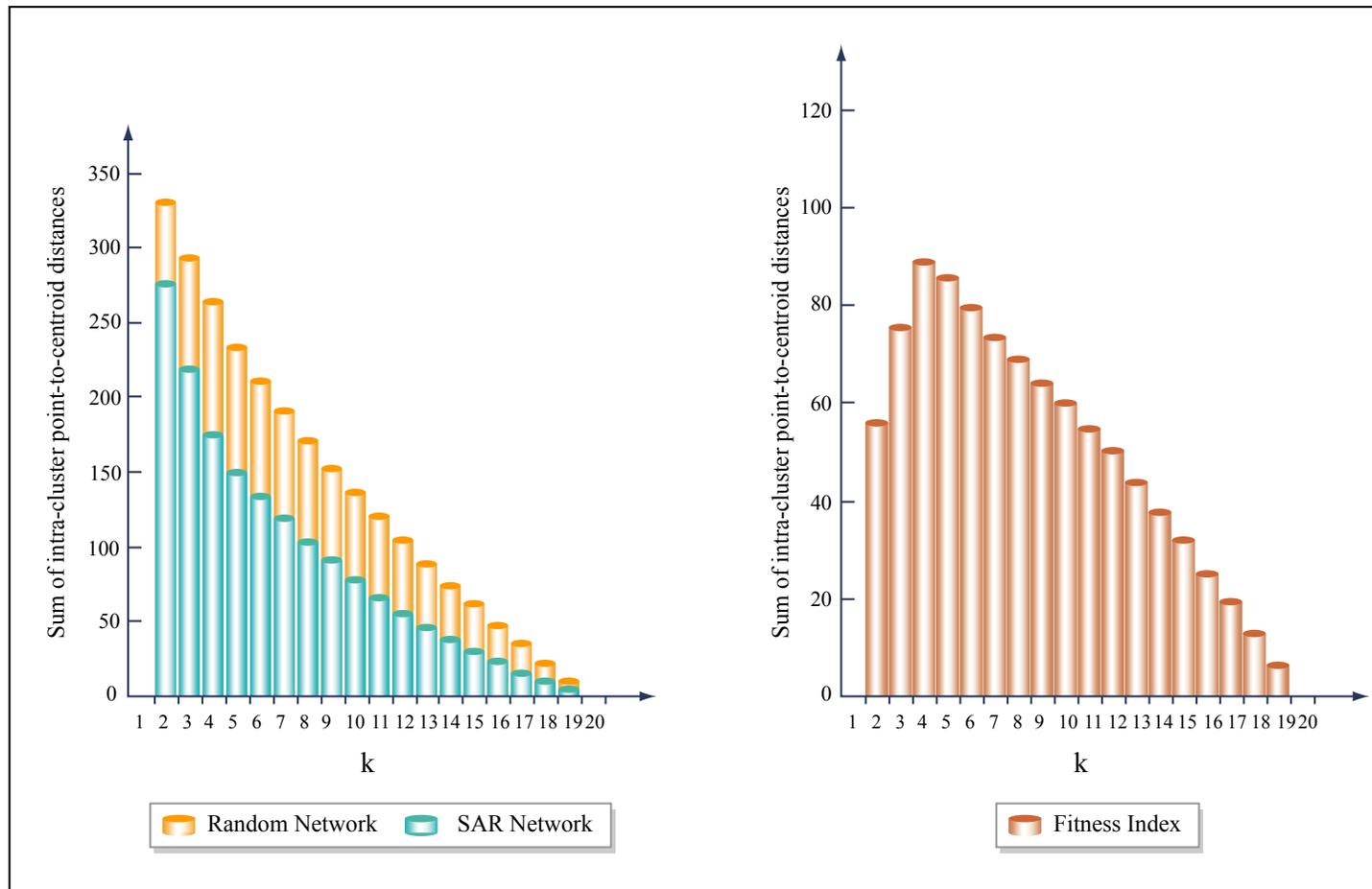

Image by MIT OpenCourseWare.

# Hsieh-Magee (Normalized) Decomposability for Structural and REGE Equivalence

□ Transform *n by n* Adjacency matrix into *n by n* Similarity matrix (using the definition of structural or REGE equivalence)

□ The sum of the inter-cluster point-to-centroid distances

$$D_k = \sum_{i=1}^{k} \sum_{j \in S_i} \left\| x_j - c_i \right\|^2$$

- ■ x_j - the *n* dimensional coordinate of node *j*
- ■ S_i *(i=1,2,...,k)* - the sub-group and
- ■ c_i - the centroid or mean point of all of the data points x_j in cluster S_i.

$$Q = 1 - \frac{D_k - D_{ideal}}{D_{max(n,k)} - D_{ideal}} = 1 - \frac{D_k}{D_{max(n,k)}}$$
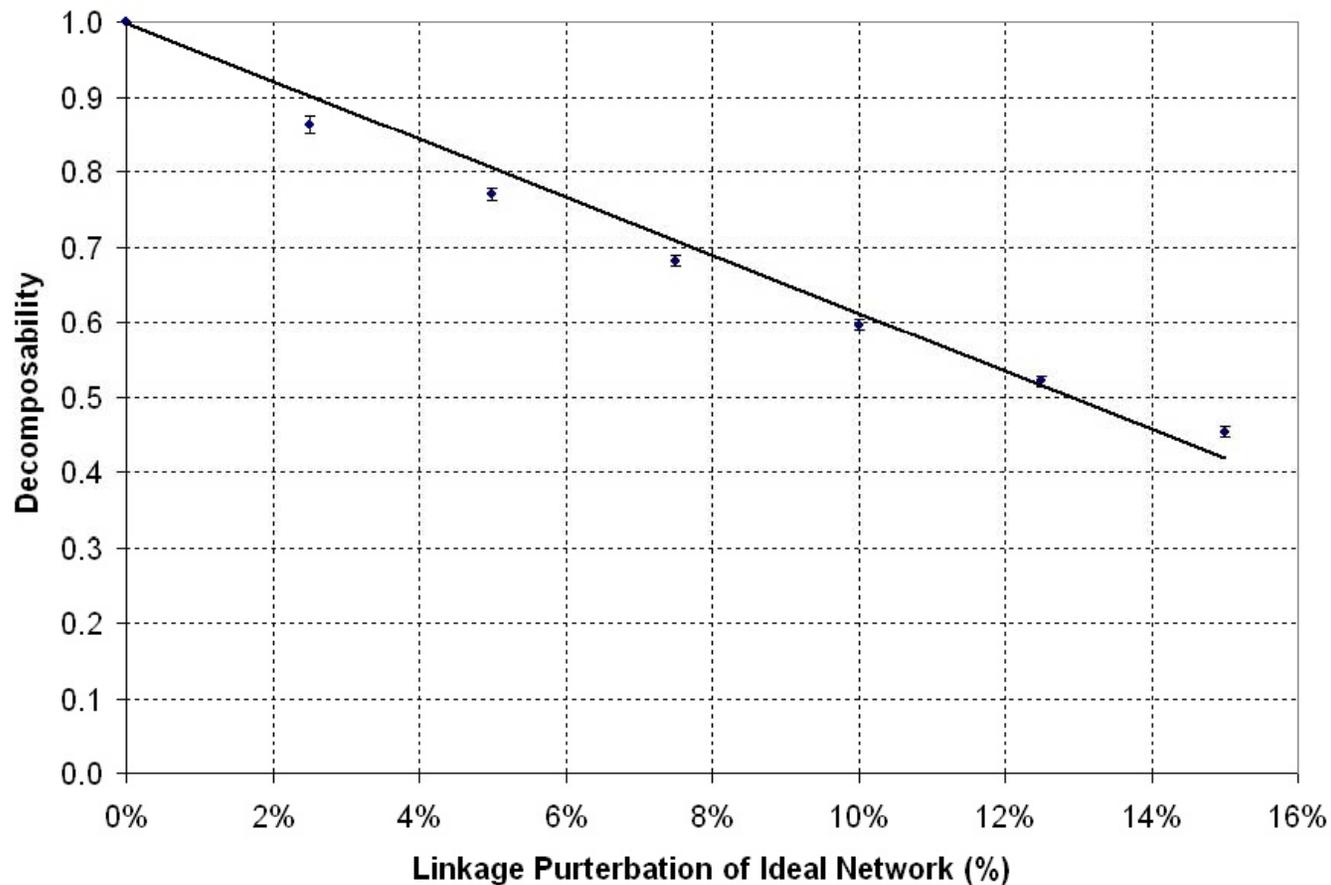
Decomposability for Regular Equivalence (II)

- ☐ REGE algorithm for measuring regular equivalence (White and Reitz 1985)
  - ■ Restrictive to only directed and acyclic networks.
  - ■ Iterative procedure in which estimates of the degree of regular equivalence between pairs of nodes are adjusted in light of the equivalence of the nodes adjacent to and from members of the pair

Decomposability for Regular Equivalence) (III)

☐ The regular equivalence for two nodes, *i* and *j*, M_{ij}, at iteration *t+1* is given by

$$M_{ij}^{t+1} = \frac{\sum_{k=1}^{N} \max_{m=1}^{N} M_{km}^{t} ({}_{ij}M_{km}^{t} + {}_{ji}M_{km}^{t})}{\sum_{k=1}^{N} \max_{m}^{*} ({}_{ij}Max_{km}^{t} + {}_{ji}Max_{km}^{t})}$$

☐ *N* - number of nodes
$max_m{}^*$ means choose the *m* according to the choice of *k* in the numerator
${}_{ij}M_{km} = min(x_{ik}, x_{jm}) + min(x_{ki}, x_{mj})$
${}_{ij}Max_{km} = max(x_{ik}, x_{jm}) + max(x_{ki}, x_{mj})$
x_{ij} - the value of the tie from *i* to *j*

Two *strategically* different approaches for decomposition of a network

□ First *quantitatively* pursued in Mo-Han Hsieh's thesis with application to decomposing the citation network of the Internet standards into meaningful subgroups but the basic ideas were developed by social network researchers 35 years ago

- Cohesion to others in subgroup
- Role similar to others in subgroup (hierarchy)
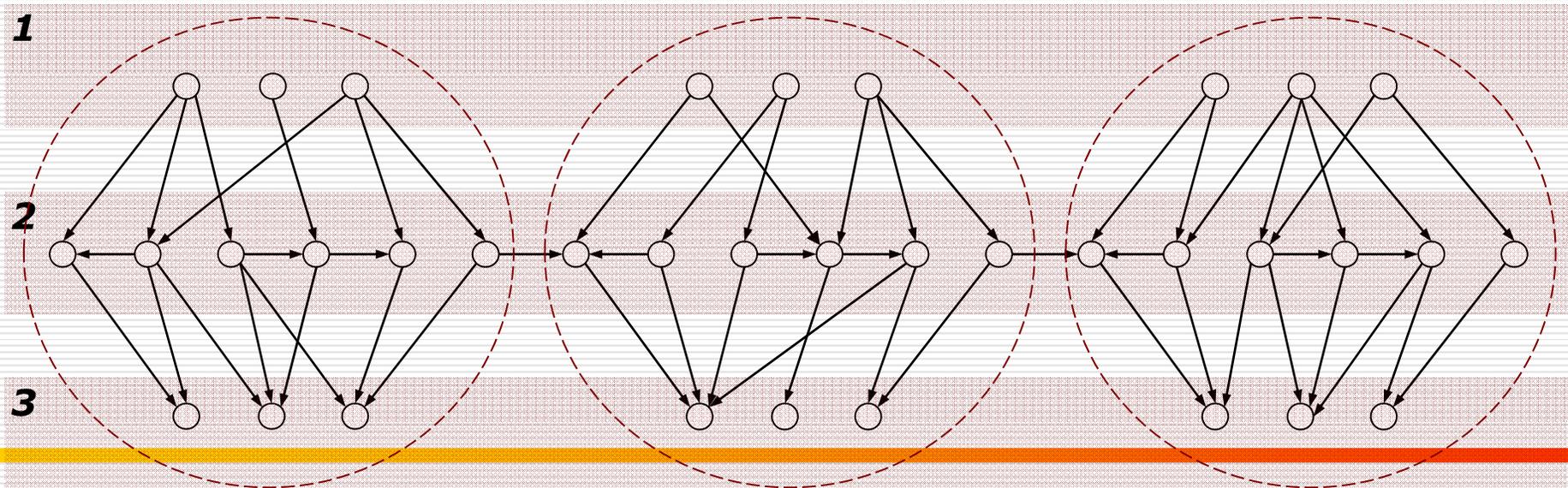
Motifs

- ☐ Milo et al. first extended the concept beyond sociological networks in a 2002 article in *Science* titled: "Network Motifs: Simple building blocks of Complex Networks",
  - ■ They defined motifs in this paper as ***patterns of interactions that occur at significantly higher rates in an actual network than in randomized networks*** and developed an algorithm for extracting them from (directed) networks

Motifs d

- ☐ Milo et al. first extended the concept in a 2002 article in *Science* titled: "Network Motifs: Simple building blocks of Complex Networks",
  - ■ They define motifs as ***patterns of interactions that are significantly higher than in randomized networks***
  - ■ They studied 19 networks (in six different classes)
    - ☐ For 2 gene transcription networks they found that the two different transcription systems showed the same motifs
    - ☐ For 8 electronic circuits (in 2 classes), they found reproducible motifs at high concentration for each class of circuit studied
- ☐ One interesting conclusion is that the technique can be applied to networks with variable nodes and links
- ☐ A second interesting conclusion coming from comparison of neurons, genes, food webs and electronic circuits is
- ☐ "**Information processing** seems to give rise to **significantly different structures than does energy flow**." The possible relevance to past Whitney work is intriguing and addressing it would involve a research question

Motifs e

- ☐ In the software tools and example data section of the web site, you can find "mfinder manual"
  - ■ This entry has links to mfinder which is software (free download) for detecting motifs on networks (PC, Windows XP and Linux versions available)
  - ■ Also comes with mDraw which allows visualization of results of mfinder.
  - ■ Also contains network randomization methods
- ☐ Biological, electronic (and social networks) have been found to have motifs and in many cases, the motifs have been valuable in understanding such systems.
- ☐ Why might electronic and biological networks in particular show motifs? What factors or *constraints* are important in these systems?

Motifs examined hierarchically

- The motifs shown for the electronic circuits (and the biological systems) seem to show evidence of functionality imbedded within the network and pursuing a hierarchy of function within technological networks is one interesting avenue suggested by this work.

Coarse-Graining

- Itzkovitz et. al. investigate Coarse-Graining as an objective means for "reverse-engineering" that can be applied even when the lower level functional units are unknown (biological focus).

- The coarse-grained version of a network is a new network with fewer elements. This is achieved by replacing some of the original nodes by CGUs (patterns of node interactions at the level being examined-motifs chosen somewhat differently).

- Itzkovitz et. al. apply simulated annealing to arrive at an optimum set of CGUs (minimize the "vocabulary" of CGUs and the complexity of the chosen CGU's while maximizing the coverage of the original network by the coarse-grained description).

Optimal selection of CGUs

☐ Complexity defined (number of "ports"for a node -equivalent to JM)

$$H = I + O + 2M$$

☐ The number of ports in the network (system) covered by a motif group selected

$$\Delta P = P_{covered} - \sum_{i=1}^{N} n_i H_i$$

☐ A scoring function which can be maximized to optimize coverage and favors CGUs which have high coverage and many internal nodes (and few external mixed nodes) is

$$S = [E_{covered} + \alpha P_{covered}] - [\alpha \sum_{i=1}^{N} n_i H_i + \beta N + \gamma \sum_{i=1}^{N} T_i]($$

Coarse-Graining b

- Itzkovitz et. al. investigate Coarse-Graining as an objective means for "reverse-engineering" that can be applied even when the lower level functional units are unknown (biological focus).

- The coarse-grained version of a network is a new network with fewer elements. This is achieved by replacing some of the original nodes by CGUs (patterns of node interactions at the level being examined-motifs chosen somewhat differently).

- Itzkovitz et. al. apply simulated annealing to arrive at an optimum set of CGUs (minimize the "vocabulary" of CGUs and the complexity of the chosen CGU's while maximizing the coverage of the original network by the coarse-grained description).

- Applying this algorithm to an electronic circuit..

Coarse-Graining c

- ☐ Itzkovitz et. al. investigate Coarse-Graining as an objective means for "reverse-engineering" that can be applied even when the lower level functional units are unknown (biological focus).

- ☐ The coarse-grained version of a network is a new network with fewer elements. This ¡s achieved by replacing some of nodes by CGUs (patterns of node interactions at the level being examined.

- ☐ Itzkovitz et. al. apply simulated annealing to arrive at an optimum set of CGUs (minimize the "vocabulary" of CGUs while maximizing the coverage of the original network by the coarse-grained description).

- ☐ Applying this algorithm to an electronic circuit, one finds a four level description which has variable functional significance and self-dissimilarity at each level

Coarse-Graining and Motifs as decomposition

☐ The *fundamental differences* between Coarse-Graining and algorithms for detection of cohesive or role equivalence community structure:

- Cohesive community structure algorithms try to optimally divide networks into **sub-graphs with minimal interconnections** but these sub-graphs are distinct and complex

- Role equivalence community structure tries to optimally divide networks into **groups of individual nodes that have similar roles** (rank in hierarchy)

- Coarse-Graining seeks a **small dictionary** of simple sub-graph types in order to elucidate the **function** of the network in terms of **recurring building blocks**

☐ It is likely that all three can give valuable information about the nature of the overall system for various networks and even simultaneously for a given network

ESD.342 Network Representations of Complex Engineering Systems

Spring 2010