DELIS

Dynamically Evolving, Large-scale Information Systems

SP 5: Biologically Inspired Techniques for “Organic IT”

Report for months 13 - 24

Participants

UniBO, UPF, Telenor, TILS

Lead partner: Bologna (UniBO)

Information Society Technologies

SP5 Goals

Goals of SP5 “Biologically Inspired Techniques for Organic IT”

Long term

Identify, understand and reverse engineer techniques inspired by biological and social systems that display “self-*” properties. Deploy these in networked information systems

Short term

Consolidate and import BISON findings. Identify “nice” properties of biological and social systems. Relate found natural network “forms” to engineering “functions”

Identify desirable life-like properties - “Self-*”

Algorithms	Simulations / Tools	Implementations

Industrial Applications

Structure of SP5 “Biologically Inspired Techniques for Organic IT”

	Months									
WP	13-15	16-18	19-21	22-24	25-27	28-30	31-33	34-36	37-39	40-42
5.1	Bio-inspired metaphors ●									
5.2	Evolved tinkering and degeneracy as engineering concepts ● ● ●									
5.3			Bio-inspired design for dynamic solution spaces ● *new partner TILS ●							
5.4	Multi-scale topology evolution in natural and artificial networks ● ● *funding reallocation (>MP)									
5.5	Identifying and promoting industrial applications and knowledge transfer ● ●									
5.6		The structure of tinkered landscapes ● *funding reallocation (new WP) ●								

● = deliverable

SP5 Deliverables Overview

Deliverables Done (by month 24)

D5.1.1: Desirable lifelike properties in large-scale information systems (month 24)
D5.2.1: Algorithms to Identify Locally Efficient Sub-graphs in Info Nets (month 12)
D5.2.2: Optimal Strategies for Construction of Efficient Info-Processing Webs (month 24)
D5.2.3: Degeneracy and Redundancy in human-constructed info. systems (month 24)
D5.4.1: Application of Motif Analysis to Artificial Evolving Networks (month 24)
D5.6.1: Classification of info.nets. - topology & functional structures & fitness landscape (month 24)

Deliverables Plan (months 25-42)

D5.2.4: Modelling open source developement networks (month 36)
D5.3.1: From biological and social algorithms to engineering solutions (month 30)
D5.3.2: Applications of bio- and socio-inspired algorithms in info. Systems (month 42)
D5.4.2: Understanding and engineering ``multi-scale'' selection in evol.nets (month 36)
D5.5.1: Promising industrial applications in dynamically evolving networks (month 30)
D5.5.2: Identifying industrial applications, examples, lessons and prospects (month 42)
D5.6.2: Integrated package for evolutionary dynamics of information networks including evolved design and landscape structure (month 36)

Goals (Start Month 13)

Long term

Inform designs for algorithms and models with direct application to network engineering and design

Short term

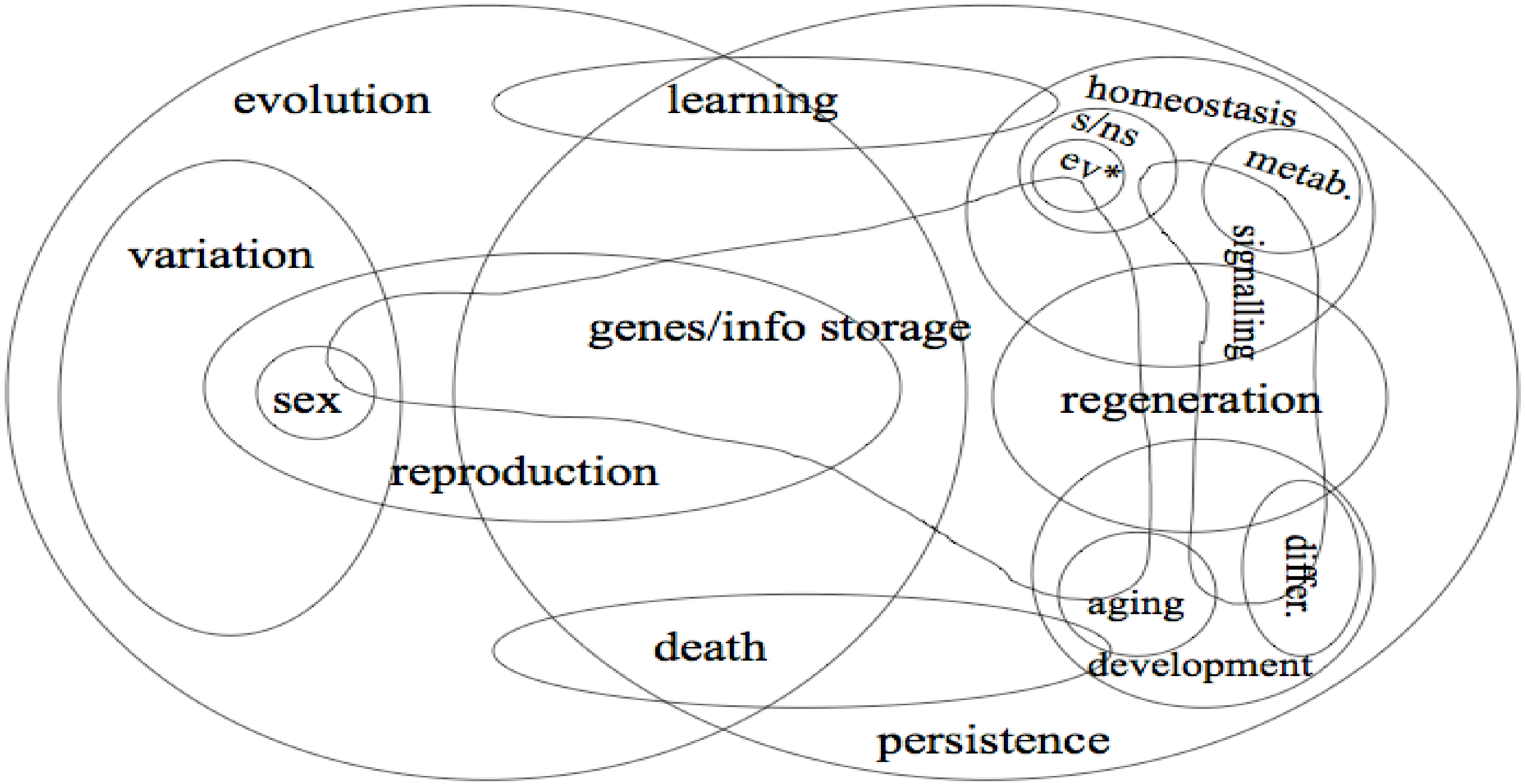
Identify a set of desirable, life-like properties in large-scale engineering systems. Review existing biologically inspired work.

Partners

Telenor, UniBO, UPF

Results (from D5.1.1)

- Identified and related desirable lifelike properties in info. systems
- Incorporated experience from concluding BISON project
- Both bio- and socio-related properties reviewed
- Some general organizational principles:
    - Modularity, Hierarchy, Self-Organization
- Some general properties for success:
    - Adaptation, Robustness, Scalability
- Also reviewed possible undesirable and problematic properties
- Related to possible application in large scale info. Systems


evolution
learning
homeostasis
s/ns
ev*
metab.
variation
signalling
genes/info storage
sex
regeneration
reproduction
differ.
aging
death
development
persistence

WP5.1 Summary

- Has relevence for all other work within SP5 and beyond
- Highly readable review and overview (D5.1.1)
- Publications:
  - Edmonds, B and Hales, D. (2005) Computational Simulation as Theoretical Experiment. Journal of Mathematical Sociology 29(3):209-232
  - Babaoglu, O. et al. (2005) "Design Patterns from Biology for Distributed Computing". Proceedings, European Conference on Complex Systems, 2005 (ECCS05). [BISON publication]
  - Márk Jelasity, Alberto Montresor, and Ozalp Babaoglu. Gossip-based aggregation in large dynamic networks.ACM Trans. Comput. Syst., 23(1):219-252, 2005 [Joint BISON / DELIS].

Goals (Start Month 0)

Long term

Explore ways of applying evolutionary computational streatgies to the optimisation of pre-existing information systems. Facilitate the interaction between engineers and automatic systems in the construction of efficient information processing networks

Short term

Investiage the topological evolution of found natural networks over time. Characterise these patterns algorithmically. Relate them to desirable functional properties for artificial engineered networks.

Partners

UPF, UniBO

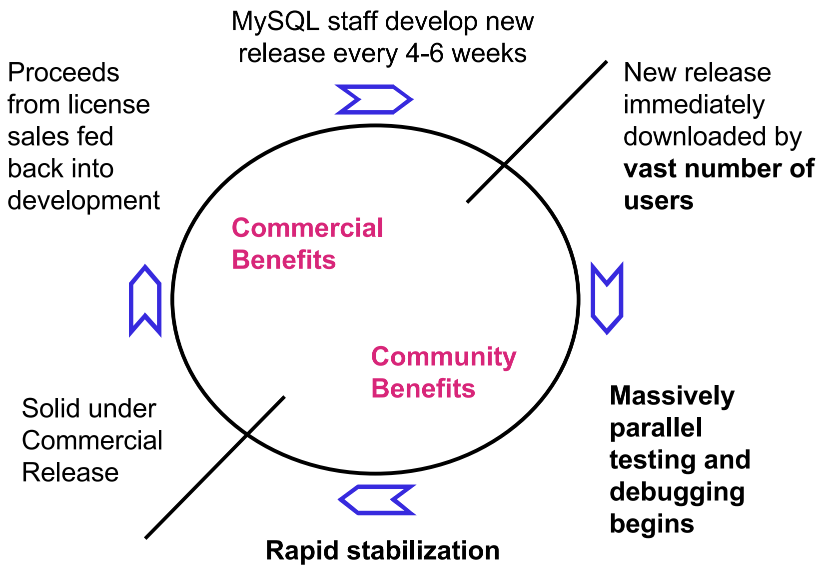
Results (from D5.2.2, D5.2.3)

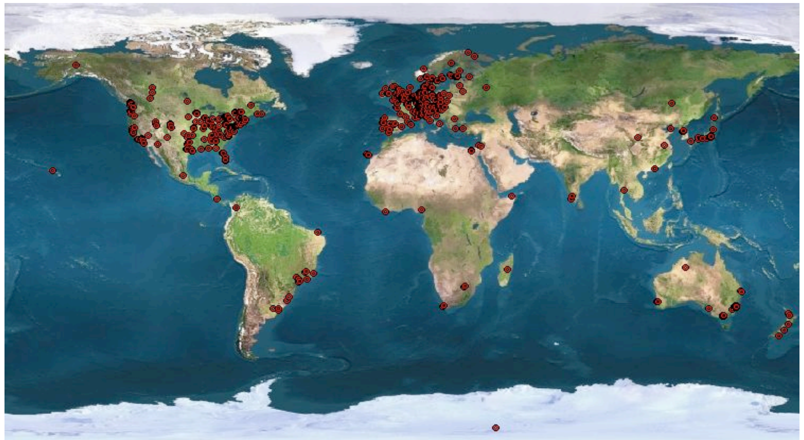
- Analysis of open source development
  - Recovery of Affiliation Networks relating developers to code
  - View open source development as co-evolution of both:
    - Programmer social networks
    - Code network represented at various scales
  - Relate to recent work on programmer social network dynamics
    - recovered from electronic discussion logs
    - Agent based social simulation models
- Exploration, analysis of relationship between tinkering, redundancy and degeneracy in evolved electronic circuits

Optimal Strategies for the Collective Construction of Efficient Information Processing Webs

What mechanisms yield successful open source projects?

Example: MySql Virtuous Development Cycle



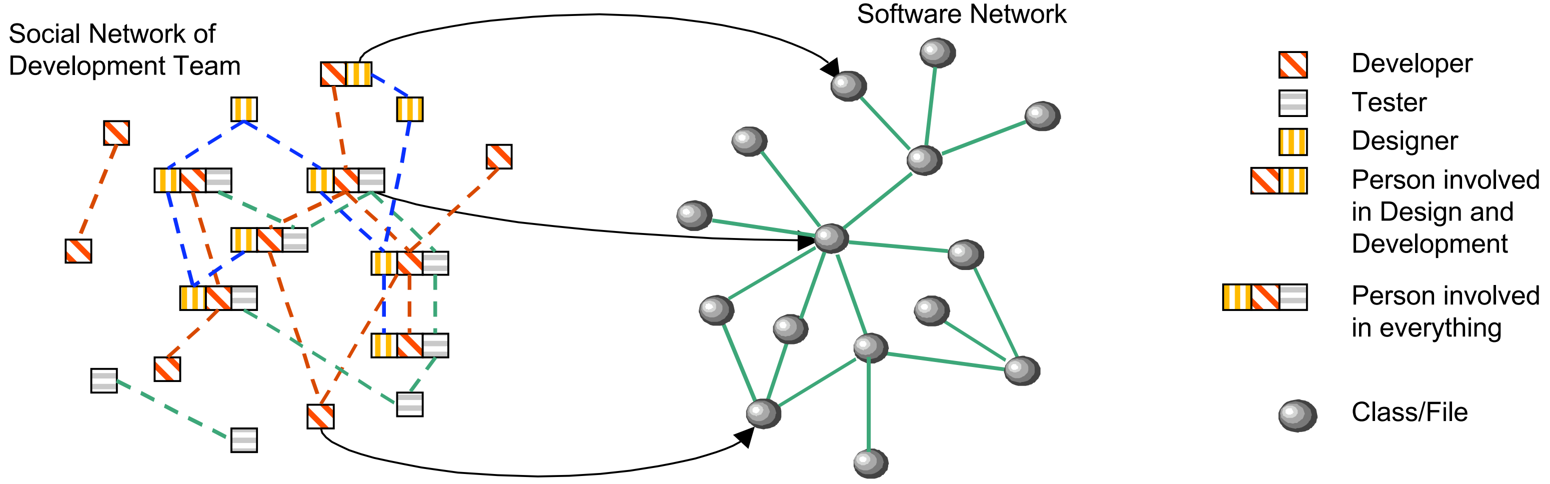


- Rapid removal of bugs
- Free of Charge
- Worldwide Distributed Development

(from http://www.debian.org/devel/developers.loc)

Optimal Strategies for the Collective Construction of Efficient Information Processing Webs

***Affiliation Networks**: What is the relationship between social networks and software networks?*

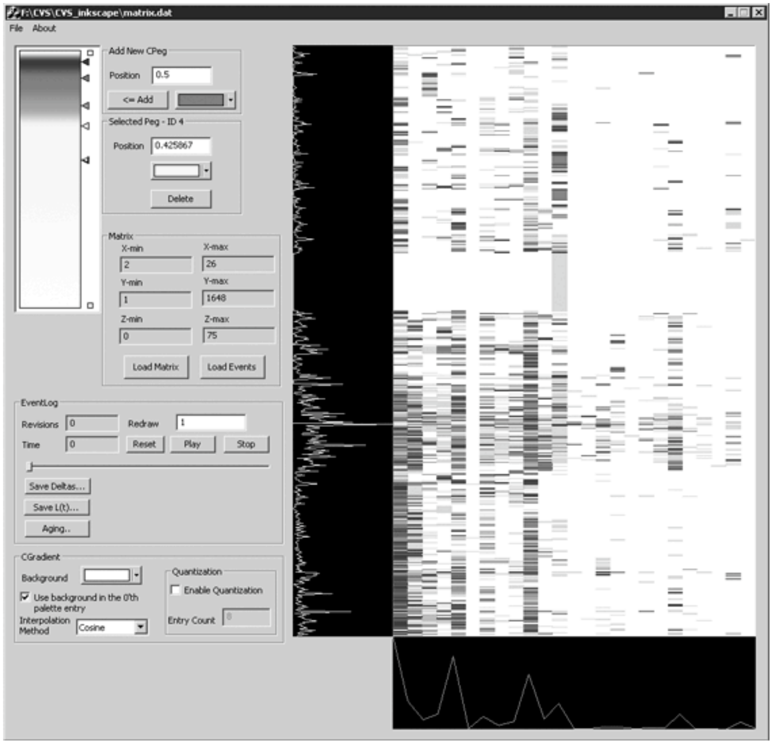


"The mapping between the Social Network of people and the Small World Network of the Software. Only a part of the entire assignment of tasks is shown with an indication of graph isomorphism and one-to-one mapping"

"Social Network Perspective of Conway's Law "
C. Amrit, J. Hillegersberg, K. Kumar,
CSCW'04 Workshop on Social Networks,
Chicago, IL, USA (2004)

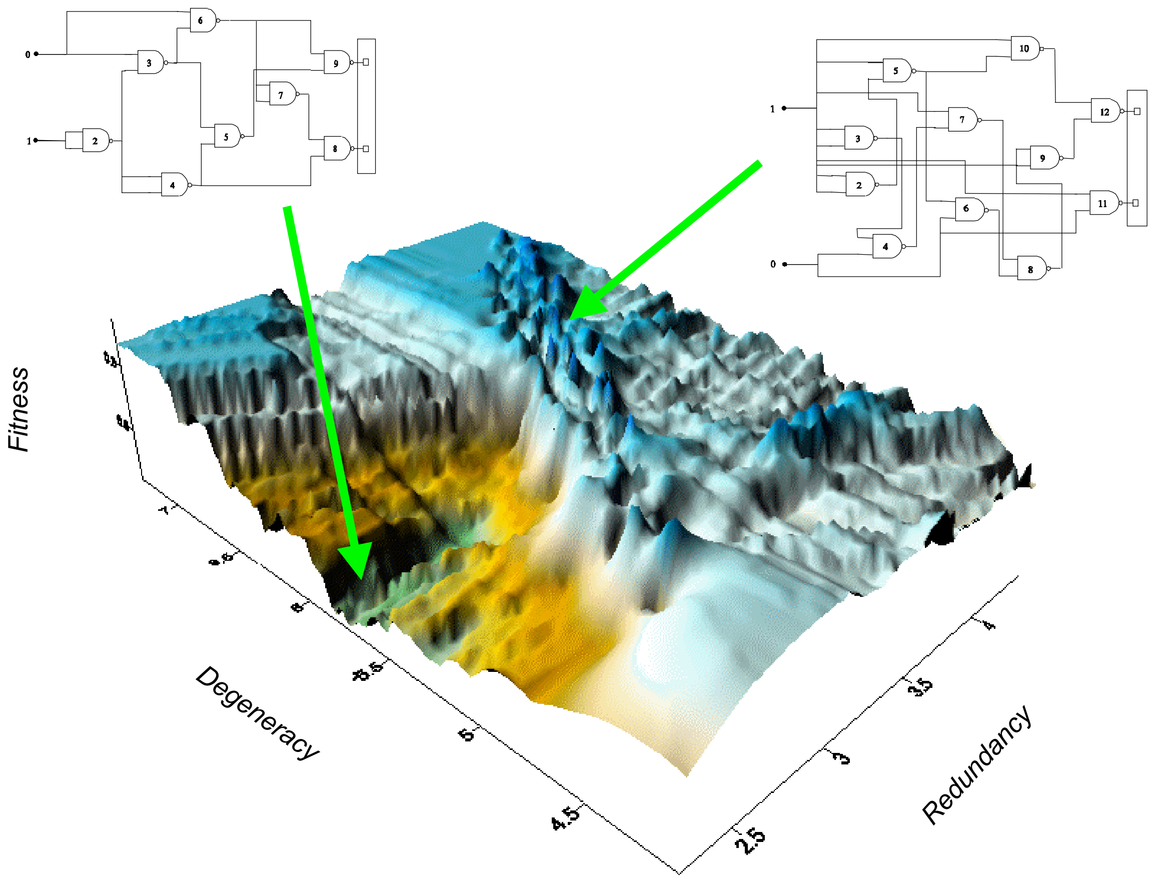
Optimal Strategies for the Collective Construction of Efficient Information Processing Webs

Software tool for recovering affiliation networks from CVS logfiles



Project	Revisions	Developers	Files
Apache	43698	78	1279
Mozilla	452101	546	28086
FreeBSD	363333	425	28056
OpenBSD	245470	195	33998
XFree86	27710	21	1788
Inkscape	15423	25	1648
SDCC	9557	32	1318
Gaim	20047	30	767
DCPlusPlus	5260	1	187

Degeneracy and redundancy in human-constructed Information systems



- Populations of digital circuits are evolved by single, random architectural changes
- Different fitness functions are used as selection criteria: we searched for maximal robustness under the presence of noise (reliability)
- Evolved robust circuits spontaneously display high degrees of degeneracy

WP5.2 Summary

- Initial work on dynamics Affiliation Nets in Open Source Dev.
- Tool to reconstruct Affiliation Nets from CSV logs
- Exploration of robustness, degeneracy and redundancy in evolved circuits
- Publications:
  - None at present
- Future: Modelling open source development networks, relating degeneracy in P2P systems (D5.2.3, month 36)

Goals (Start Month 19)

Long term

Develop tools and methods to translate / modify biologically and socially inspired algorithms for application in realistic information systems environments

Short term

Select a set of candidate algorithms and application domains. Use simulation and apply necessary tuning using

Partners

UniBO, UPF, Telenor, TILS

WP5.3: Biologically and socially inspired design for dynamic solution spaces

On-going (started month 19)

- Select ideas from other SP5 WPs applicable to *realistic* distributed engineering problems
- Identify the engineering constraints / requirements that differ from the existing algorithms
- Develop tools and methods to translate / modify the algorithms
- Working on Cooperative Resource Replication model with TILS

Deliverables Planned

D5.3.1: From Biological and social algorithms to engineering solutions (month 30)
D5.3.2: Applications of bio- and socio-inspired algorithms in info. Systems (month 42)

WP5.4 Multi-Scale topology evolution in natural and artificial networks

Goals (Start Month 13)

Long term

Explore processes of general network evolution in both natural and artificial systems - determine and harness both the form and function of multi-level evolution for engineering

Short term

Apply “motif analysis” to artificial networks developed for functional properties and compare with natural systems with similar or desired properties. Relate network forms to functions.

Partners

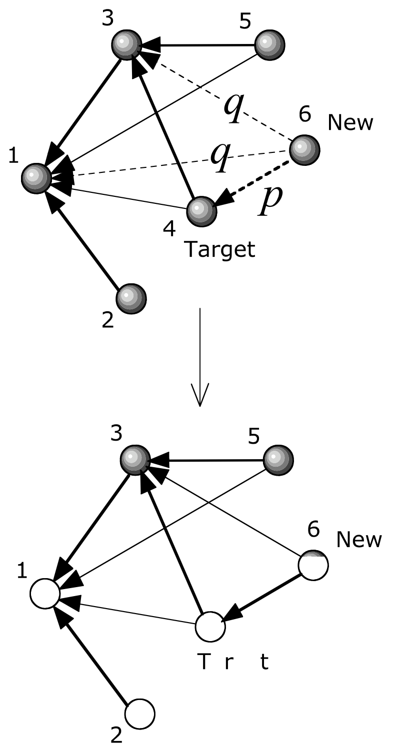
UPF, UniBO, Telenor

Results (from D5.4.1)

- Evolution of software code networks
  - Based on the assumption that software code networks evolve by a copy and re-wire process (not related to function)
  - Model of evolution of structure of software nets
  - Produces predictions that match data from software dev. logs.
- Motif analysis of evolving P2P networks
  - Application of motif analysis to two developed P2P protocols
  - Protocol SLAC (see D5.2.1) uses simple copy and re-wire rule to emerge and sustain cooperation between nodes
  - Protocol SLACER, a probabilistic modification of SLAC producing cooperative and connected networks

Application of Motif Analysis to Artificial Evolving Networks

Growing Network with Copying (GNC) model



$$\frac{dL}{dN} = mp + mq\frac{L}{N}$$

Evolution of number of links L(t)

$$P_i(k_i) \approx k^{-2}$$

Scale-free in-degree distribution
(<u>independent</u> of copying parameters)

$$\frac{dL}{dt} = \left[mp + mq\frac{L}{N(t)} \right] \dot{N}(t)^{-1}$$

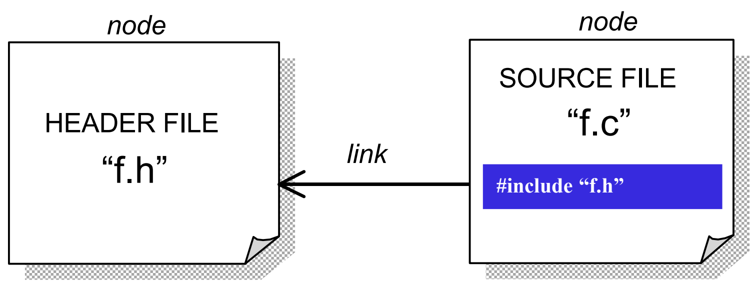
Time-dependent evolution

"Network Growth by Copying"
P.L. Krapivsky and S. Redner,
Physical Review E, 71, 036118 (2005)

"Logarithmic Growth Dynamics of Software Networks" **S. Valverde and R. V. Solé,**
Europhysics Letters 72 (5) pp. 858-864 (2005)

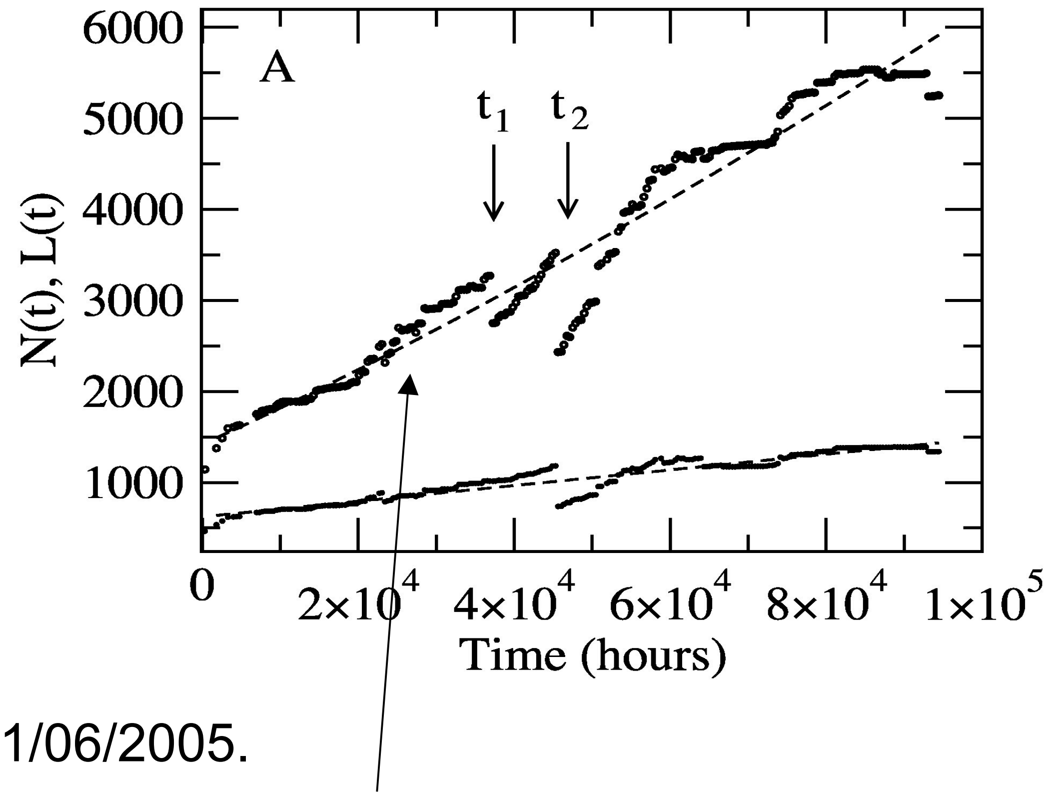
WP5.4 Evolution of Software Nets

Example: First prediction of number of #include's in a C/C++ project



$N(t) = N_h(t) + N_c(t)$ = Number of project files

$L(t)$ = Number of "#include" clauses

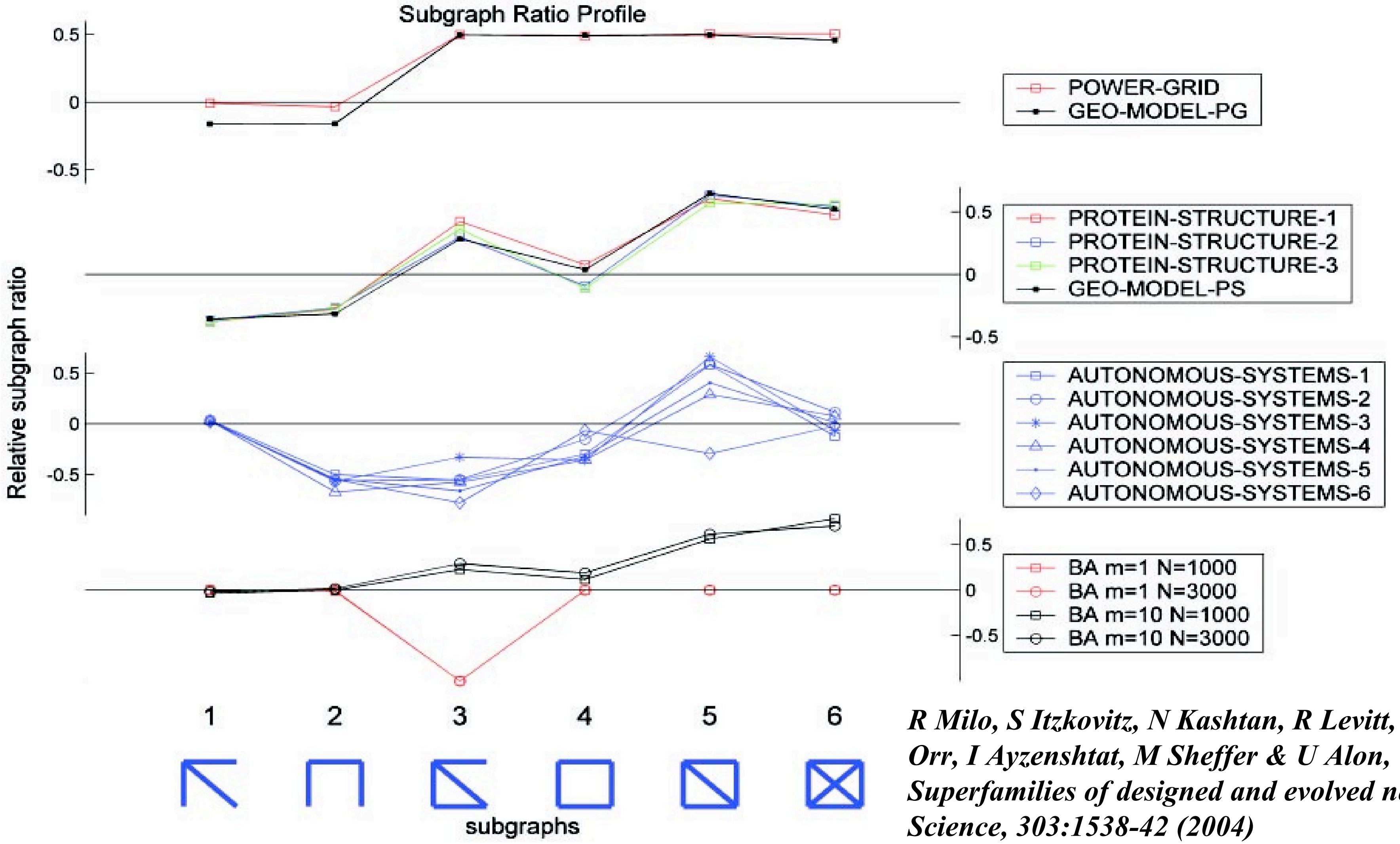


XFree86 between 16/05/1994 and 01/06/2005.

Assume linear growth of N. GNC model predicts L(t)

"Logarithmic Growth Dynamics of Software Networks" **S. Valverde and R. V. Solé,** Europhysics Letters 72 (5) 858 (2005)

WP5.4 Motifs in evolved nets



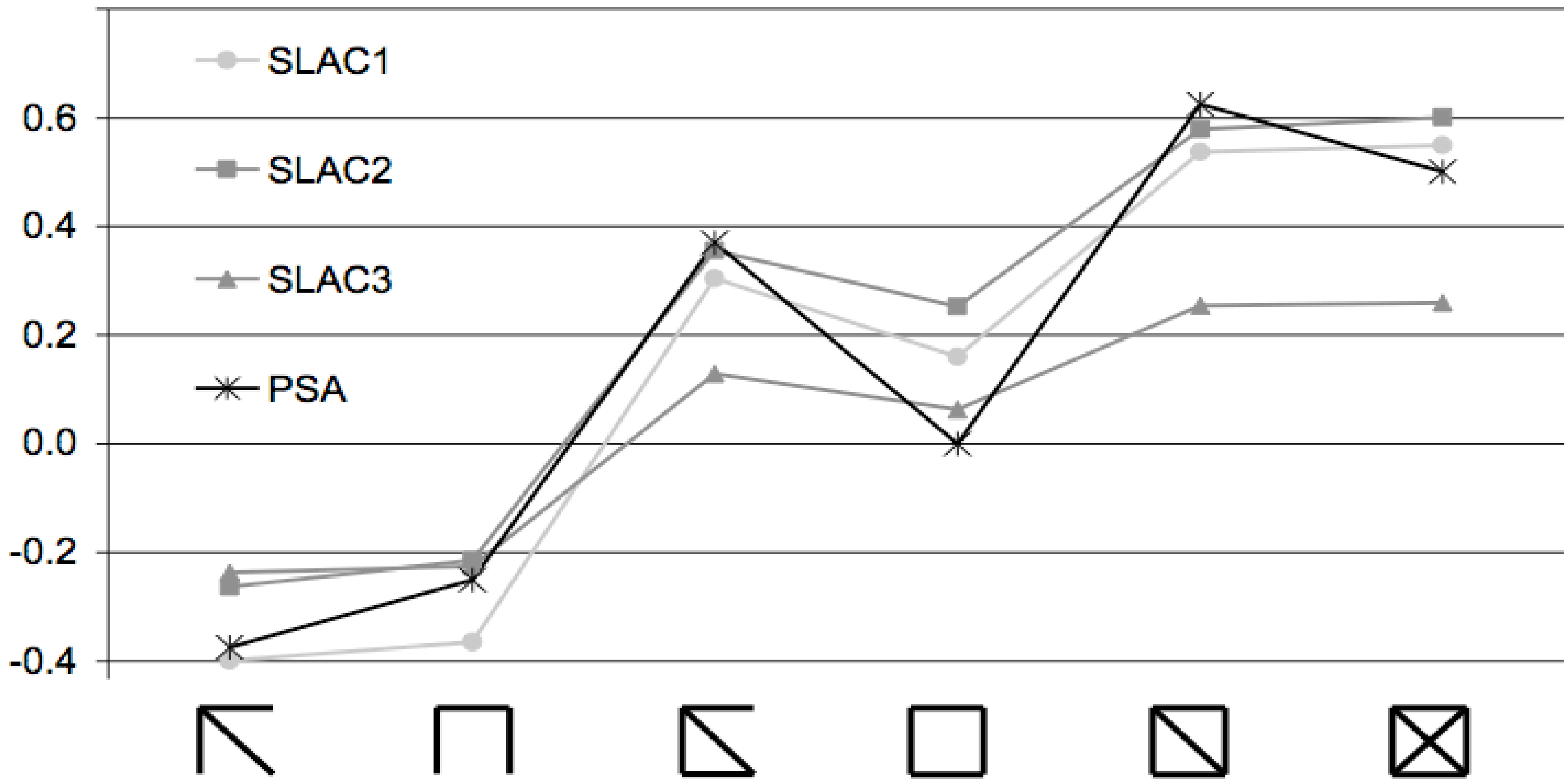
R Milo, S Itzkovitz, N Kashtan, R Levitt, S Shen-Orr, I Ayzenshtat, M Sheffer & U Alon, Superfamilies of designed and evolved networks Science, 303:1538-42 (2004)

Basic SLAC node-level algorithm

(has some “nice” properties - as previously reported see D5.2.1)

- Periodically **do**
  - Compare “utility” with a random node
  - **if** that node has higher utility
    - copy that node’s strategy and links (reproduction)
    - mutate (with a small probability):
      change strategy (behavior)
      change neighborhood (links)
  - **fi**
- **od**

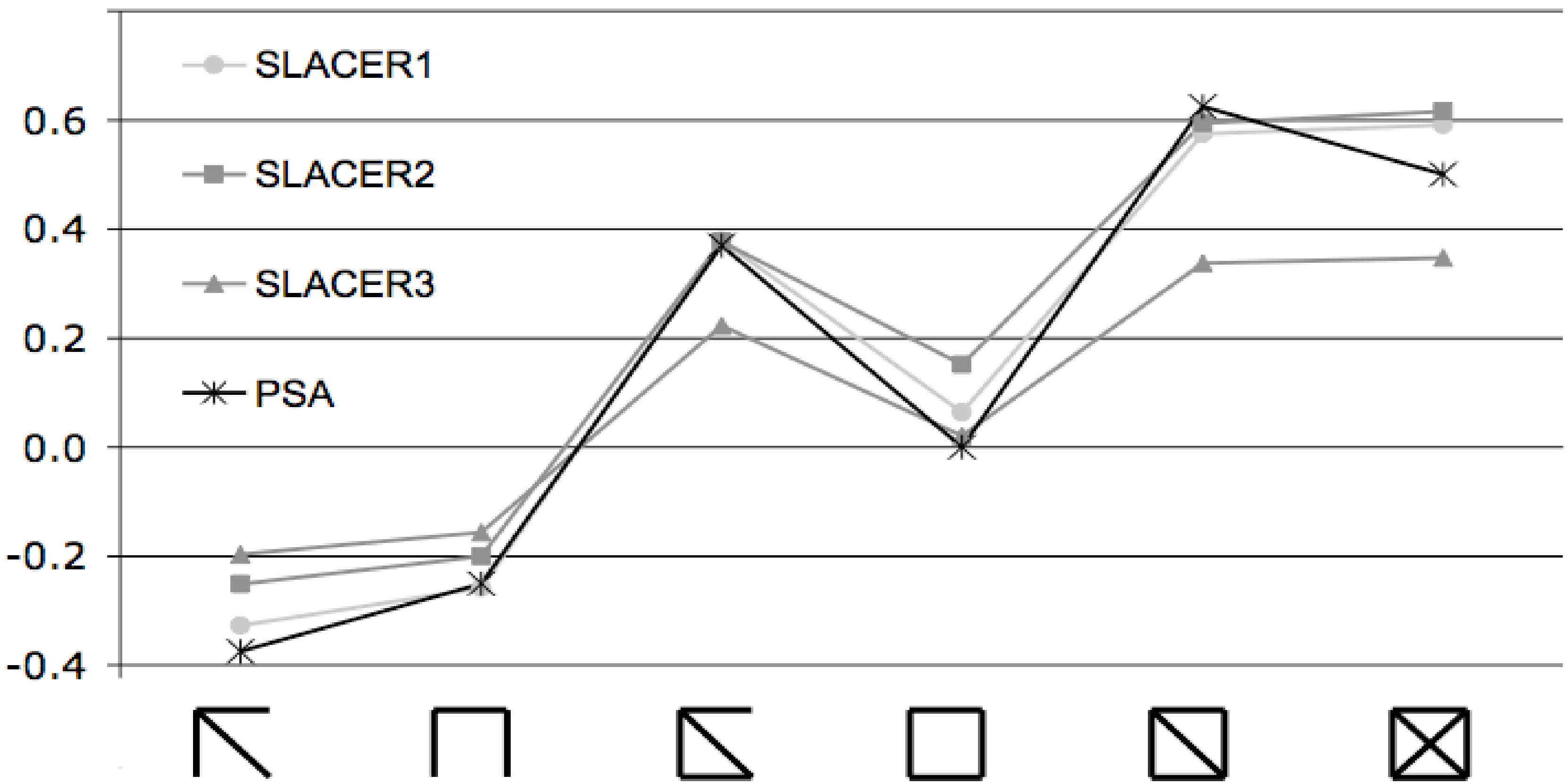
WP5.4 Motifs in evolved P2P nets



Network size N = 500, edges E ≈ 3500.

SLAC1, 2, 3 taken immediately before, during and after high cooperation breaks out.

WP5.4 Motifs in evolved P2P nets



SLACER - a probabilistic form of SLAC producing small-world type topologies.

WP5.4 Summary

- Predictive analysis of software development - potential uses in software metrics
- Motif analysis of SLAC P2P protocol - interesting links to natural systems, potential use for monitoring performance
- Publications:
  - “Logarithmic Growth Dynamics of Software Networks” S. Valverde and R. V. Solé, Europhysics Letters 72 (5) 858 (2005)
  - “SLACER: randomness to cooperation in peer-to-peer networks” Hales, D.; Arteconi, S.; Babaoglu, O. Proc. of Workshop on Stochasticity in Distributed Systems (STODIS’05), IEEE Computer Society Press (2005).
- Future: further predictive metrics, motif-based network monitoring, distributed real-time motif estimations in evolving P2P (D4.5.2, Month 36)

WP5.5 Industrial applications and knowledge transfer

Goals (Start Month 13)

Long term

Bridge between academic research (in DELIS SP5) and realities of industry (telecom). Patents, spin-offs, industrial projects

Short term

Identify SP5 activities and mechanisms with possible commercial and industrial applications

Partners

Telenor, UniBO, UPF

WP5.5 Industrial applications and knowledge transfer

On-going (started month 13)

- Number of promising areas that could be considered:
- Fully distributed power method (potential for distributed document ranking) mainly in SP6 (UniBo, Telenor)
- open source community structures - design and mangmnt. SP5 (UPF)
- motifs in software networks - software dev. & maintenance SP5 (UPF)
- cooperative P2P with healthly community structures SP5 (UniBo)

Deliverables Planned

D5.5.1: Promising industrial applications in dynamically evolving networks (month 30)
D5.5.2: Identifying industrial applications, examples, lessons and prospects (month 42)

WP5.6 The Structure of Tinkered Landscapes

Goals (Start Month 16)

Long term

Comparison of biological networks and engineered designs
Understand evolutionary mechanisms that make natural networks robust and have other differing properties. Produce simulator package.

Short term

Characterize topologies, functional constraints, fitness landscapes of existing networks. Relate knowledge to optimizing evolutionary rules / algorithms.

Partners

UPF, UniBO

Results (from D5.6.1)

- Experiments with evolved feed-forward networks and analysis of fitness landscape properties
- Some counter-intuitive insights
- Similar properties to RNA folding
- Relate to potential in P2P systems - tentative

Information networks and their fitness landscapes

RNA molecules have neutral landscapes

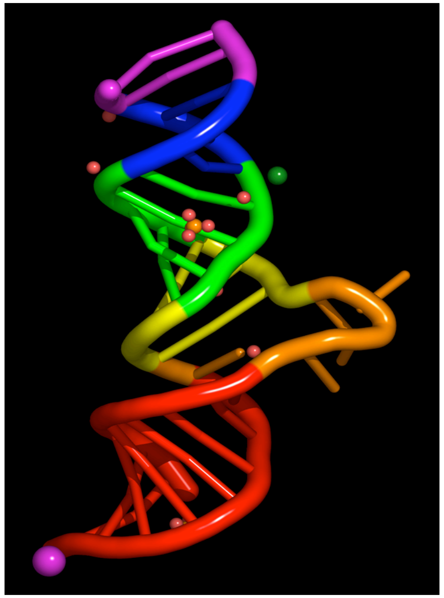
GENOTYPE = STRUCTURE

5’ **GUGAUGG...GGUUAC** 3’

RNA sequence

folding

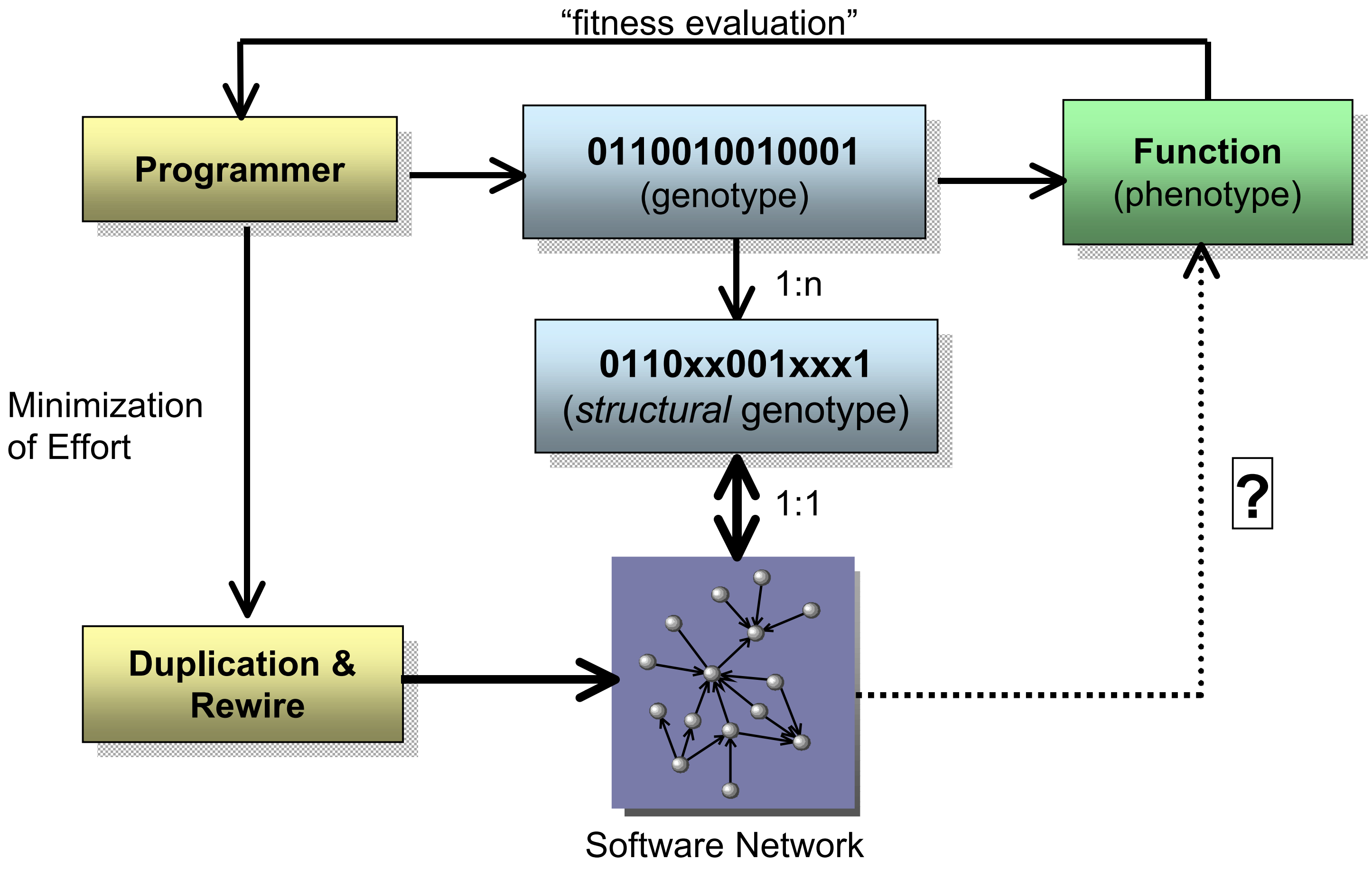
PHENOTYPE = FUNCTION



RNA shape

- Hypothesis: *the fitness landscape of networks performing information processing might help understanding how they evolve and how easily can be evolved.*

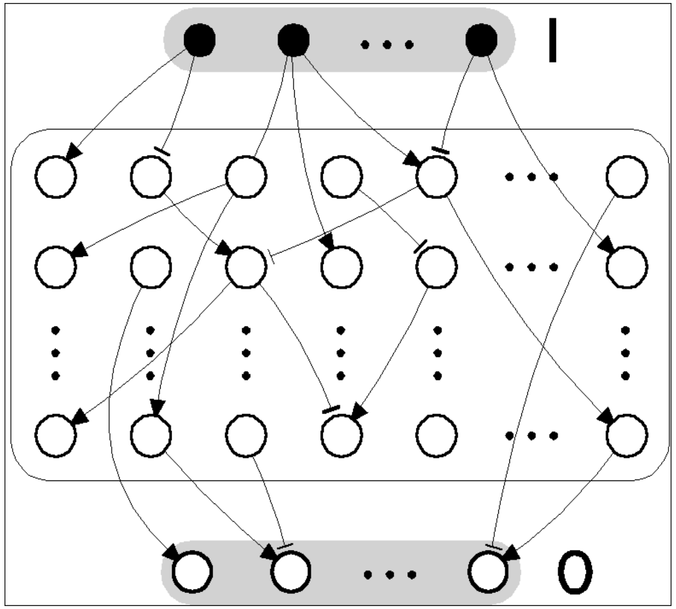
What is the landscape of software systems?



Information networks and their fitness landscapes

Case Study: building bio-inspired computational networks

Network of binary linear zero-threshold units - “perceptrons”
Outputs +1 if input threshold > 0
Weights on links: +1, 0 or -1
Genotype = ordered string of weights
Phenotype = implemented boolean function from inputs (I) to outputs (O)
Mutation = remove one link, add a new one with prob(1/3) of -1, prob(2/3) +1



Feed forward networks

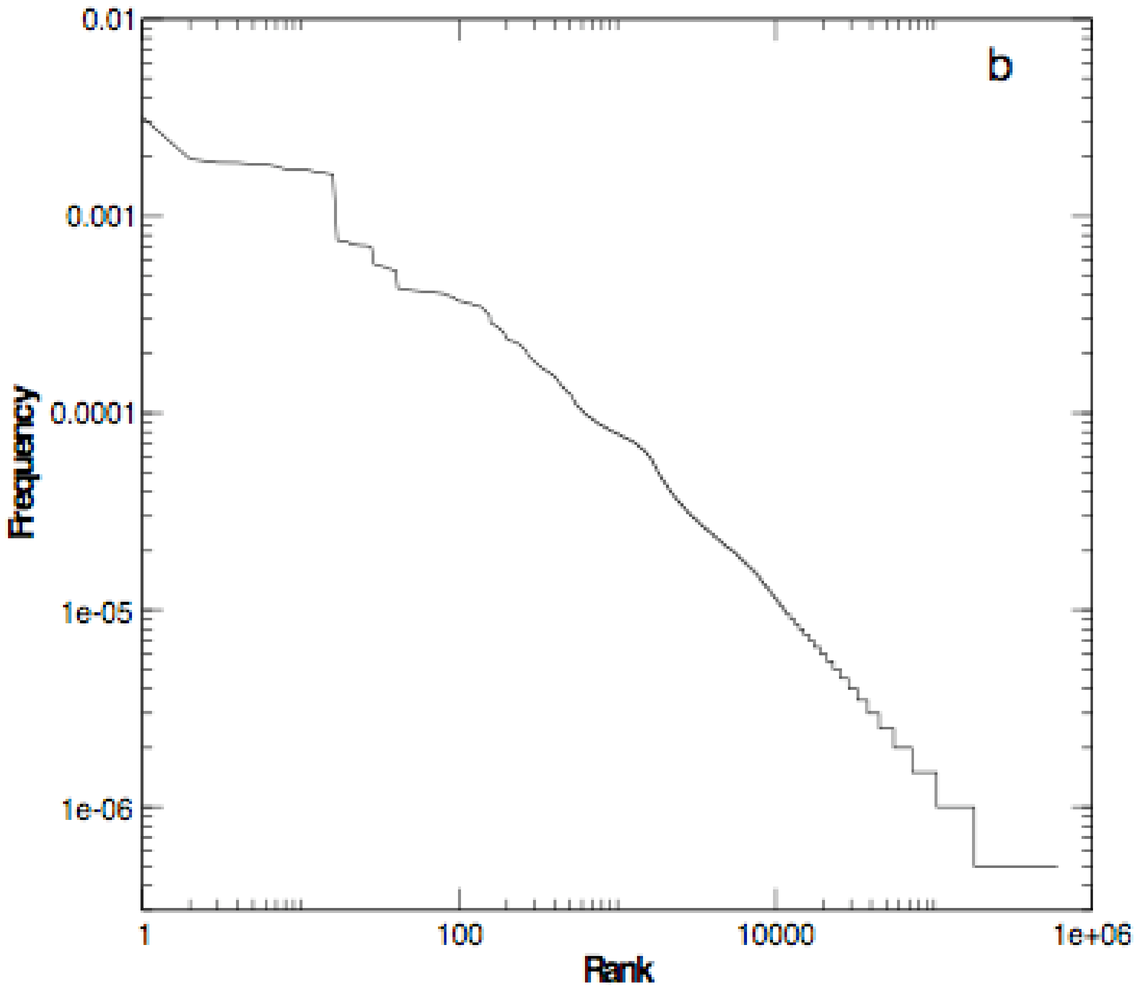
FEED-FORWARD LANDSCAPES ARE EQUIVALENT TO RNA LANDSCAPES

WP 5.6 Tinkered landscapes

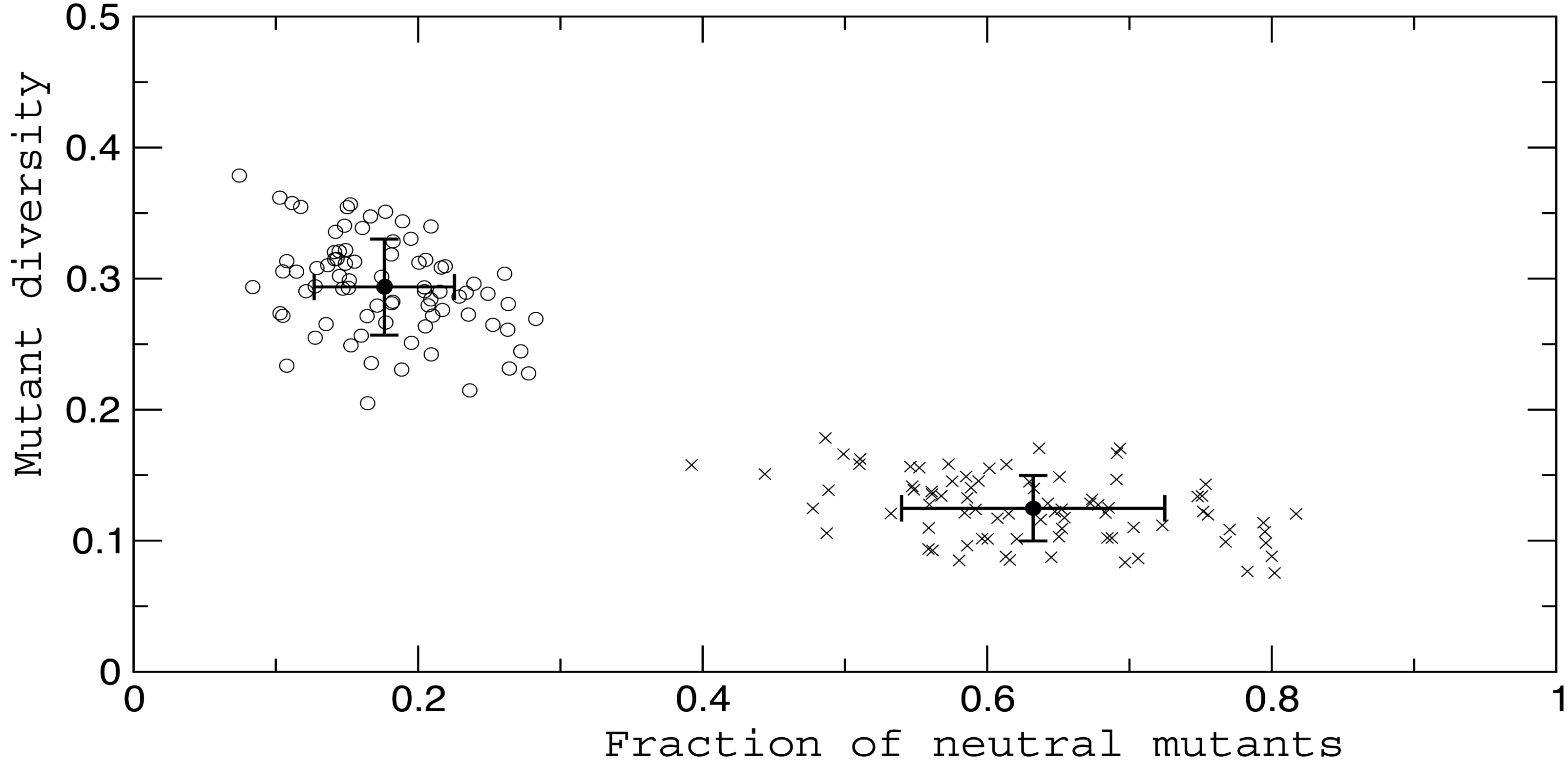
Random sample of genotypes
Many genotypes => same phenotype

Frequency of different phenotypes follows a power law (like RNA folds)

Chart shows rank-frequency of of genotypes by function (many functions common, some very rare)



WP 5.6 Tinkered landscapes



Random sample of genotypes in space (x's) shows high neutrality and low diversity of mutants. After hill-climbing for opposite in G-space while preserving function (phenotype) can find points in G-space (o's) - "portals" to many different functions (phenotypes)

WP5.6 Summary

- Feed-forward networks demonstrate many of the properties of RNA fitness landscapes => robustness but also "portals"
- In dynamic P2P link and node failure and churn can be viewed as "mutation" of the structure. The aim is robust function under these
- The protocol is the "genotype" => self-org. structure => function
- Publications:
  - Fernandez, P. and V. Sole, R. (2005) From wiring to function and back: a case study infeed-forward networks. Santa Fe Inst. Working Paper.
  - Hales, D. and Arteconi, S. (2005) Friends for Free: Self-Organizing Artificial Social Networks for Trust and Cooperation. [DELIS-TR-0196]
- Future: integrated package for exploring landscapes, potential applications to P2P design (D5.6.2, Month 36)

SP5 Dissemination and Cooperation

Cooperation with other SP's

SP4-SP5 Game theory and evolutionary economics models
SP5-SP6 Cooperative distributed information sharing
SP1-SP5 Possibility of better dynamic visualisation of P2P (planned)
CCT2, CCT3 Meetings attended

Cooperation with other projects

- BISON — As described, extensive cooperation with concluding BISON
- NANIA — EPSRC (UK) 5 year project - collaborative meetings planned / already made, with Manchester group
- CATNETS — On-going collaboration (FET STREP)
- ONCE-CS — Complexity Network, High presence at ECCS'05

Dissemination

Popular press: New Scientist (Jan 2005), Atlas Magazine (March 2005), P2Pnet and Slashdot news websites (June 2005), Business week (Dec 2005).

Thank you!