HUMBOLDT-UNIVERSITÄT ZU BERLIN

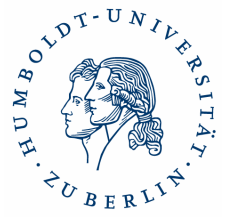
Simple is Beautiful:

a Comparison-based Diagnosis

Miroslaw Malek
Institut für Informatik
Humboldt-Universität zu Berlin
malek@informatik.hu-berlin.de

Venue: Sofitel Chicago Water Tower
Chicago
June 25, 2010

Outline



Introduction

Comparison Model (with comparators)

MM - Models

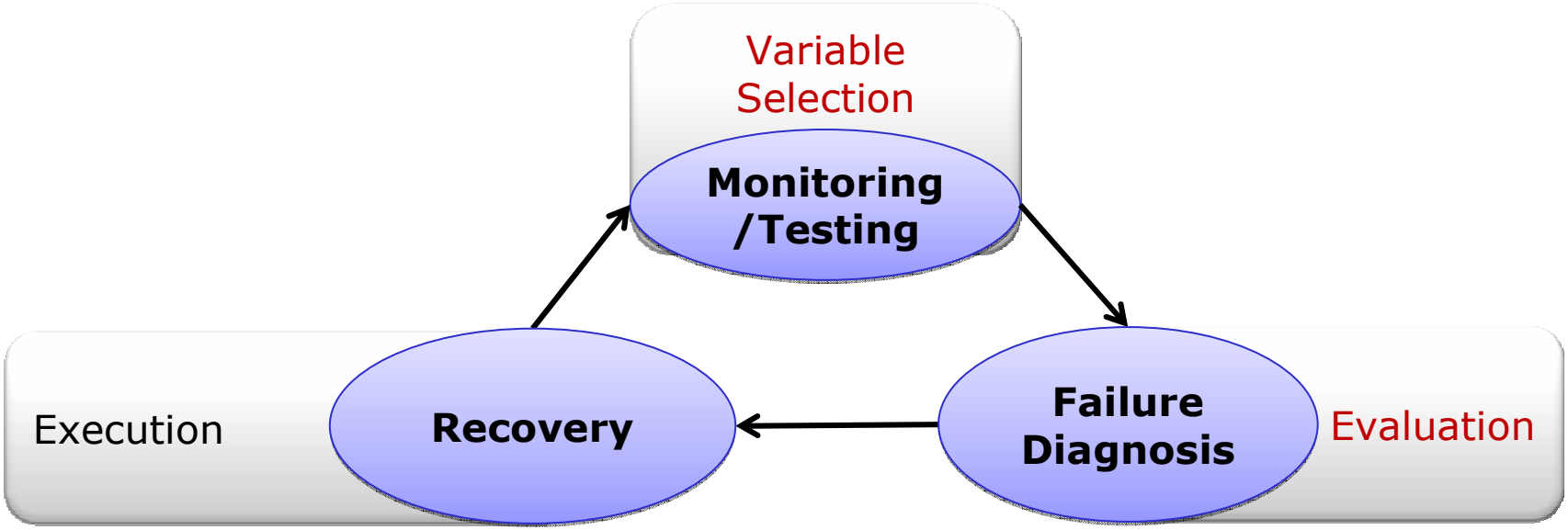
Key extensions

Implementations

Challenges

Failure Diagnosis in Cycle

Failure diagnosis is an essential part of fault management and is usually followed by recovery actions.

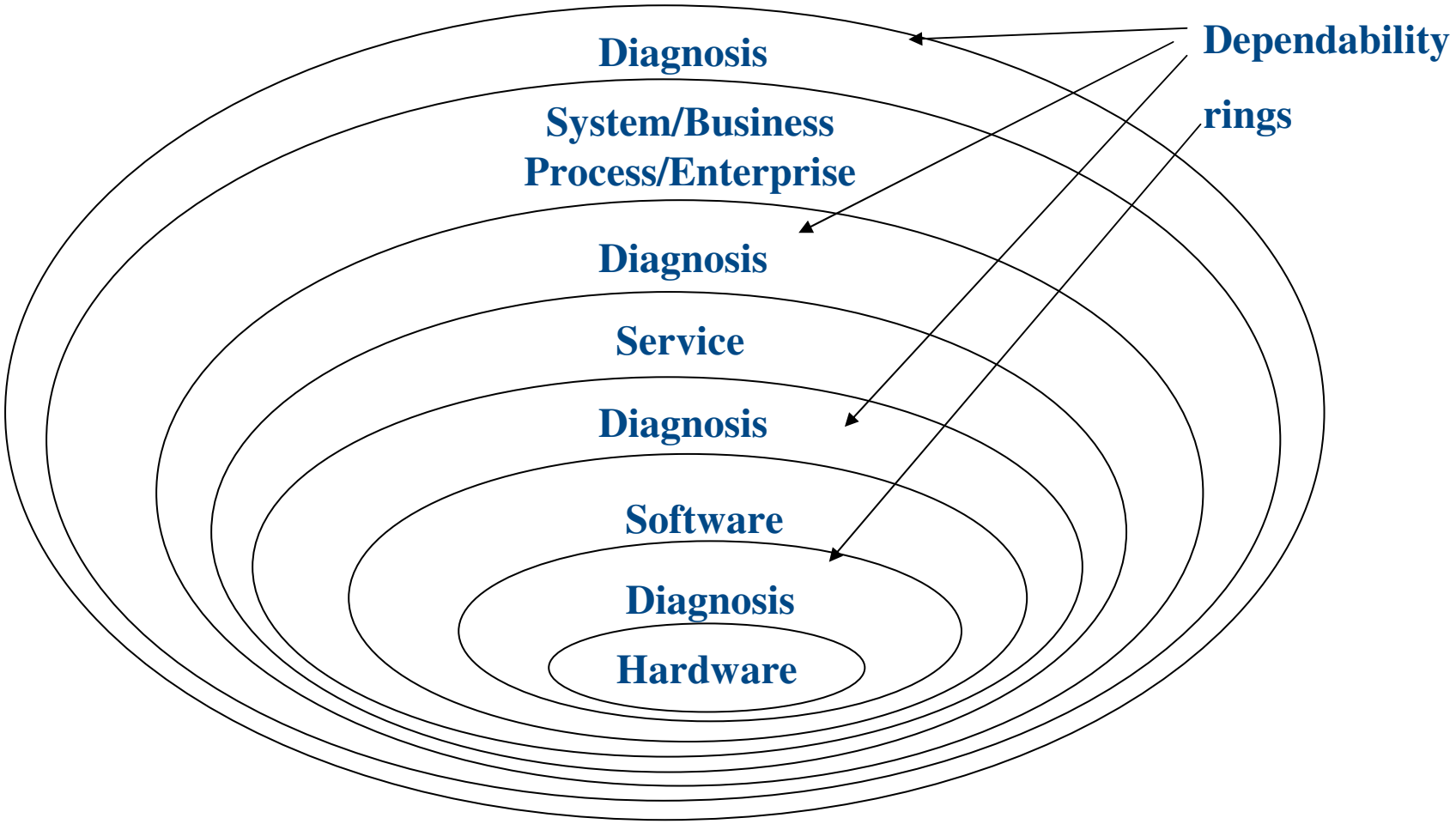


Failure Diagnosis (continued)

- Packaging, testability, diagnosability and performance instrumentation are frequently afterthoughts or are developed independently in the design process
- Use of concurrent error detection is frequently indispensable (especially in multiprocessor/cloud environments) due to high system complexity and rapid system contamination
- Diagnosis should cover all system levels

- In this talk: **Emphasis on application (algorithmic) and system level diagnosis**

Translucency – Getting the Biggest Bang for the Buck



At what level providing measures and mechanisms for diagnosis and proactive fault management will maximize the payoff (minimize downtime)?

Three Phenomena that Won't Go Away

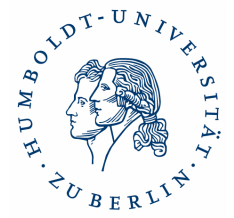
- Ever-increasing systems **complexity**
  - Growing connectivity, chip density and interoperability
  - Growing number of functionalities
- Increasing **uncertainty**
  - Ever-growing number of attacks and threats, novice users and third-party or open-source software, COTS
  - Ever new failure modes
  - Dynamicity (frequent configurations, reconfigurations, updates, upgrades and patches, ad hoc extensions)
- Increasing real-**time** requirements
  - Systems proliferation to applications in all domains of human activity where many of them require real time
  - Growing users expectations regarding timeliness

Therefore, diagnosis is and will remain a permanent challenge.

The Key Principle: KISS

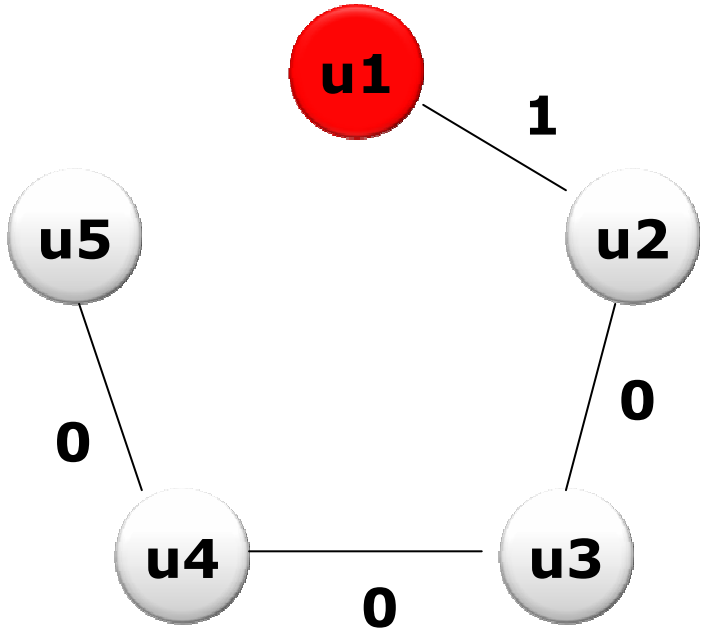
- With ever-increasing systems complexity **simplicity** is of an essence
- Striving for **simplicity** and keeping all stages of the system design and development simple is a major challenge
- Divide-and-conquer, integration, interoperability and structured design principles and hierarchical approaches should be applied to all aspects of design and maintenance. These main methods are insufficiently exploited design of various functionalities/properties such as testability, diagnosis, real time, performance monitoring, etc.
- In this talk the focus is on enforcing simplicity in system diagnosis

The Comparison



- The comparison is an essential concept from beginning of times
- In computers the comparison is widely used:
  - Password
  - Bank account, identity checking
  - Signatures, counters, results of computations
  - Testing and diagnosis, watchdogs, etc.
- First fault-tolerant systems have used comparison in duplex system for failure detection
- Examples include AT&T's ESS and 3B20 system series

Basic Comparison Models



Unit 1	Unit 2	Comparison Outcome
fault-free	fault-free	0 (pass)
fault-free	faulty	1 (fail)
faulty	fault-free	1 (fail)
faulty	faulty	1 (fail) or X

- Each edge corresponds to a comparator
- |_ (n+1)/2_| comparators (node cover) guarantee detection
- n-1 comparators are sufficient for a single node diagnosis
- n(n-1)/2 comparators assure (n-2)-diagnosability (t=n-2)

(Malek, 1980, Chwa and Hakimi, 1981)

Definition of t-Diagnosability

A system of n units is one step t-fault diagnosable (t-diagnosable) if all faulty units within the system can be located without replacement, provided the number of faulty units does not exceed t.

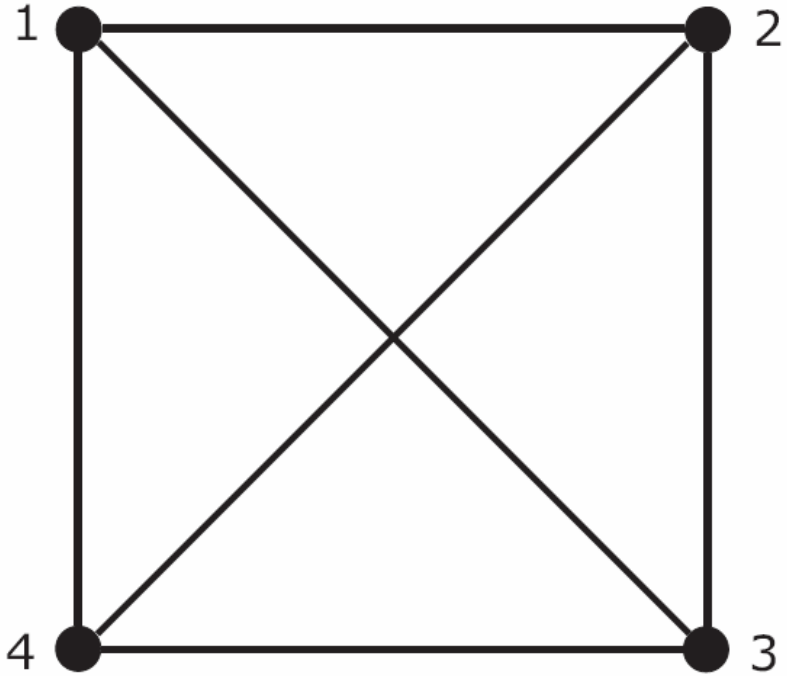
1. $2t+1 \leq n$
2. At least t units must test each unit

(Preparata-Metze-Chien)

- Several diagnosis algorithms have been proposed, with a variety of assumptions

Diagnosability in a Comparison Model

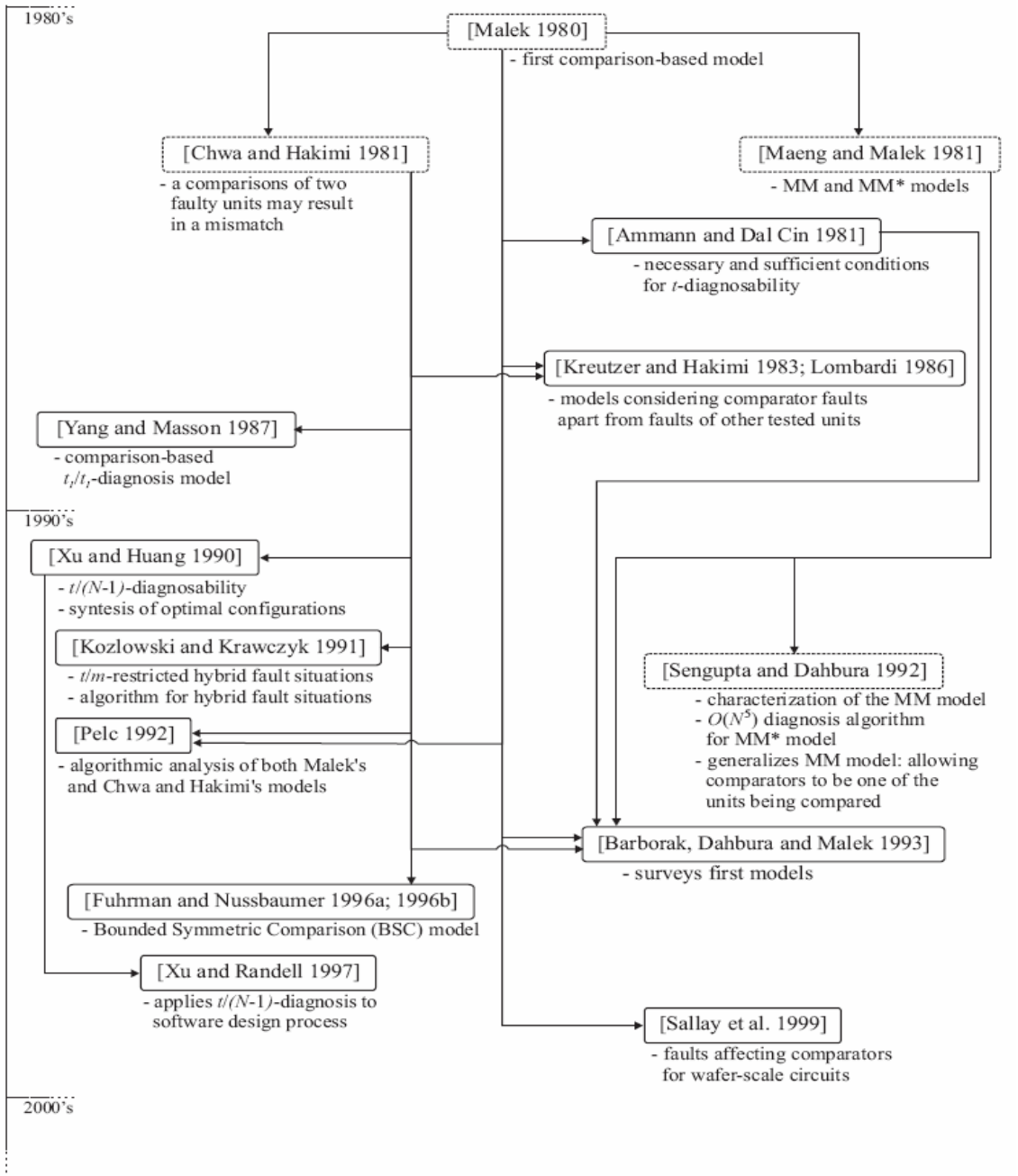
- Edges indicate comparators between pairs of units
- A complete graph is (n-2)-diagnosable for n>3
- In general, n - |_n/3_| comparators (for n > 2) are sufficient for diagnosis under a single fault assumption and up to n nodes fault detection



An example graph representing comparisons among four units

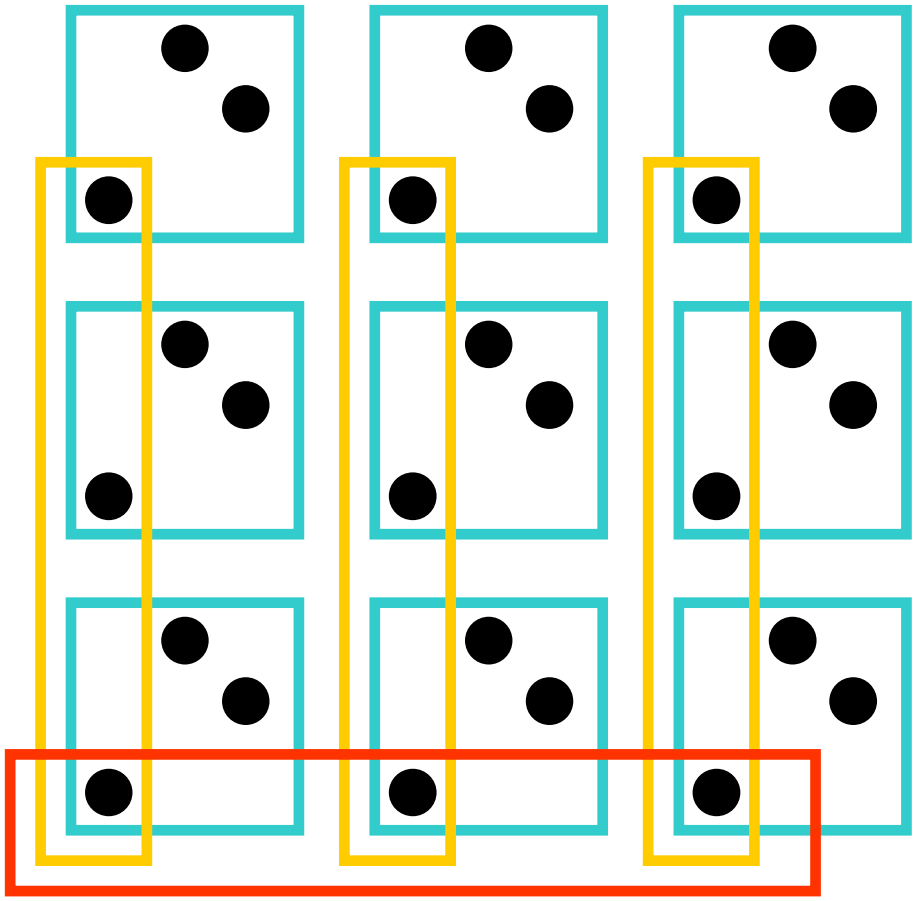
Model	Reference	Main Contributions
Malek's model	[Malek 1980]	- first comparison-based model - compared units are different - the comparison of one or two faulty units results in a mismatch - central observer is a trusted unit that executes comparisons and performs the diagnosis - the diagnosability is $N-2$
	[Ammann and Dal Cin 1981]	- necessary and sufficient conditions for t-diagnosability
	[Sallay et al. 1999]	- strategy to identify faults affecting comparators - application for wafer-scale circuits
	[Pelc 1992]	- algorithmic analysis of both Malek's and Chwa and Hakimi's models - worst case number of tests for optimal algorithms for t-diagnosis, sequential t-diagnosis and one-step t-diagnosis for both models, under non-adaptive and adaptive testing
	[Barborak et al. 1993]	- surveys early models
Chwa and Hakimi's model	[Chwa and Hakimi 1981b]	- the comparison of two faulty units may result in a match
	[Fuhrman and Nussbaumer 1996b; 1996a]	- Bounded Symmetric Comparison model, considers a limit on the number of faulty units that can produce identical results
	[Kozlowski and Krawczyk 1991]	- extension of Chwa and Hakimi's model for t/m-restricted hybrid fault situations
	[Yang and Masson 1987]	- comparison-based t_1/t_1-diagnosis model
	[Xu and Huang 1990]	- characterization of $t/(N-1)$-diagnosability under Chwa and Hakimi's model - synthesis of optimal $t/(N-1)$-diagnosable configurations for topologies such as chains and loops
	[Xu and Randell 1997]	- application of $t/(N-1)$ diagnosis to the software design process
	[Kreutzer and Hakimi 1983; Lombardi 1986]	- models considering comparator faults apart from faults of other tested units - characterization of the proposed models, $(t-t_c)$-diagnosability

Summary of comparison-based results based on early models from Duarte, Roverli, Ziwich, Albini, 2010

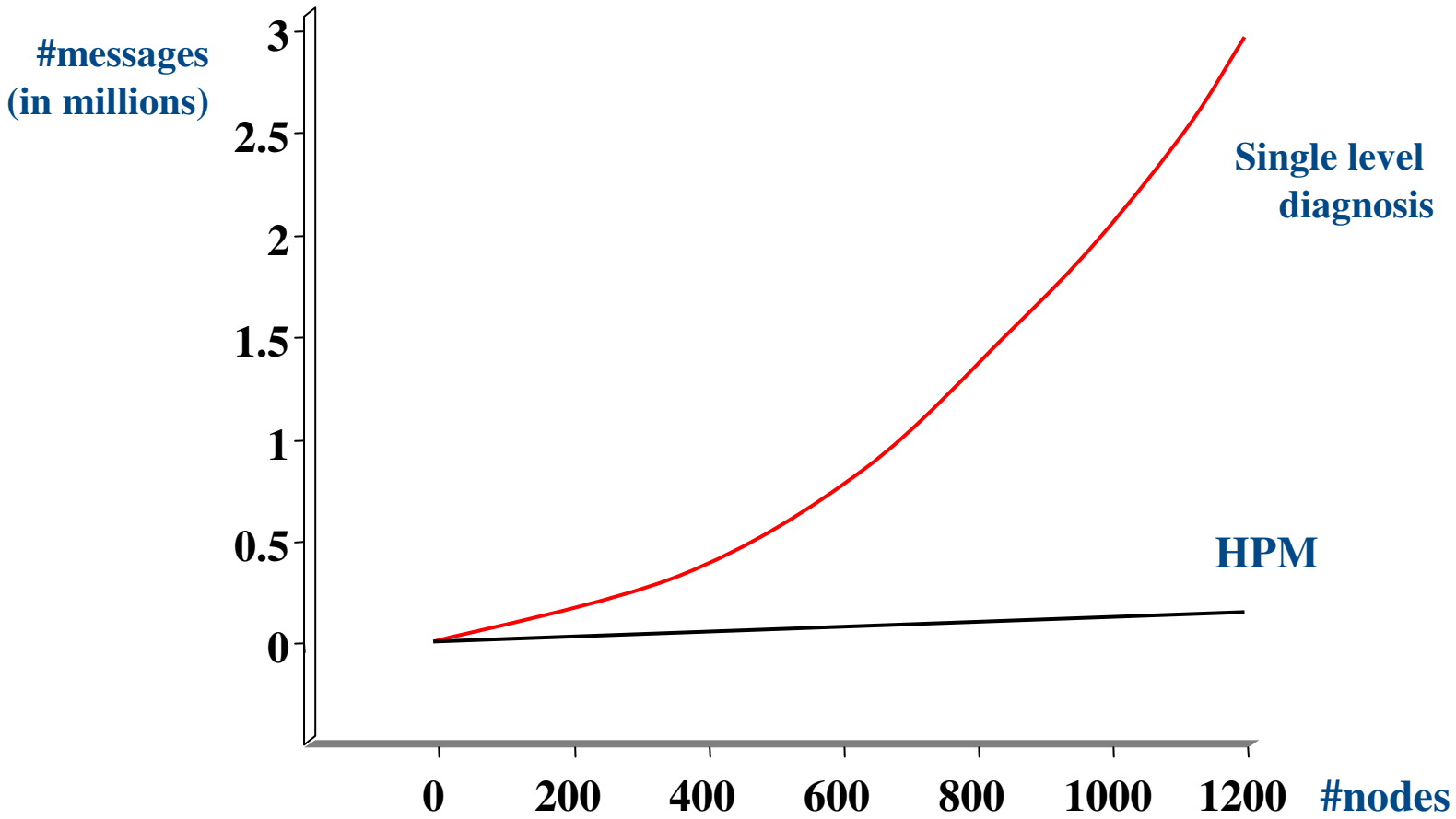


Comparison-based diagnosis timeline: results based on early models from Duarte, Roverli, Ziwich, Albini, 2010

Hierarchical Diagnosis

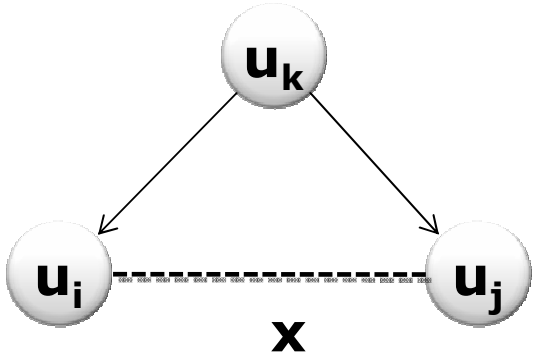


The Hierarchical Diagnosis Performance



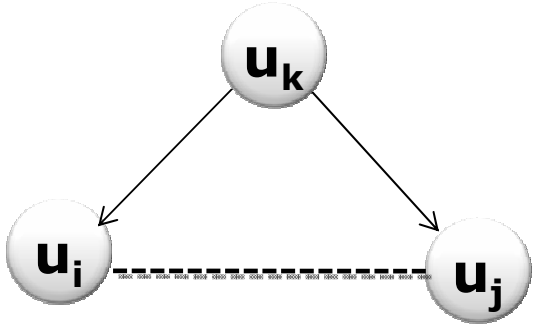
The MM-Comparison Models (1981)

- A year later we have proposed a different approach:
  - The testing processor sends some test input to two adjacent nodes or asks for signature or counter values
  - The testing processor compares the two responses and sends the outcome to the central diagnosis unit
- The comparison graph is built where two nodes u_i and u_j are connected by an edge if there is a testing node u_k that tests u_i and u_j.
- Graph theory based algorithms can be used to identify the set of faulty processors



Maeng and Malek (1981)

MM-Comparison Model



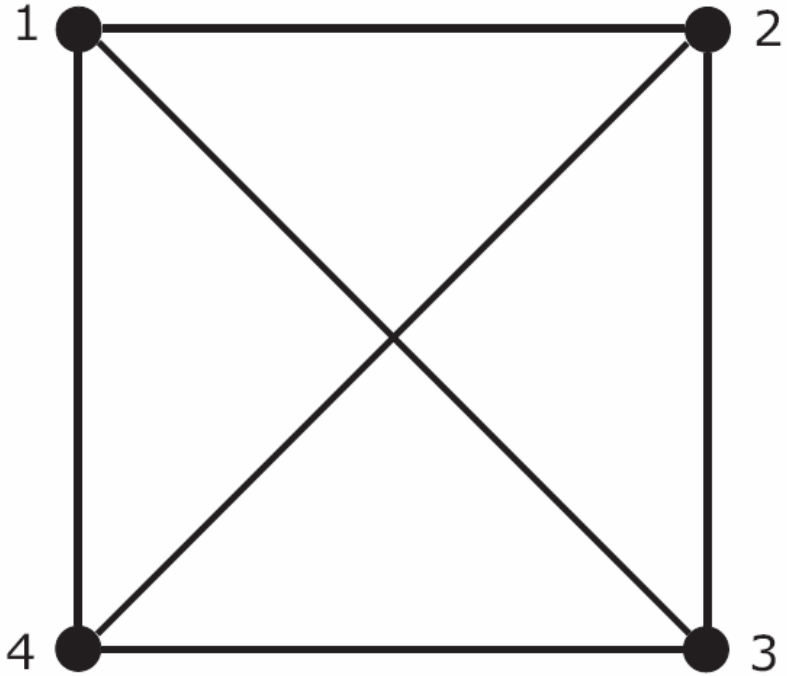
Comparator (Unit k)	Unit i	Unit j	Comparison Outcome
fault-free	fault-free	fault-free	0 (pass)
fault-free	fault-free	faulty	1 (fail)
fault-free	faulty	fault-free	1 (fail)
fault-free	faulty	faulty	1 (fail)
faulty	fault-free	fault-free	0 or 1
faulty	fault-free	faulty	0 or 1
faulty	faulty	fault-free	0 or 1
faulty	faulty	faulty	0 or 1

Diagnosability in an MM-Model

- Edges indicate connections in the system

- A graph is t-diagnosable iff d(v) >t-1 and a condition on duals to prevent ambiguities

- Also an algorithms for generating an optimal graph for t >3 has been proposed

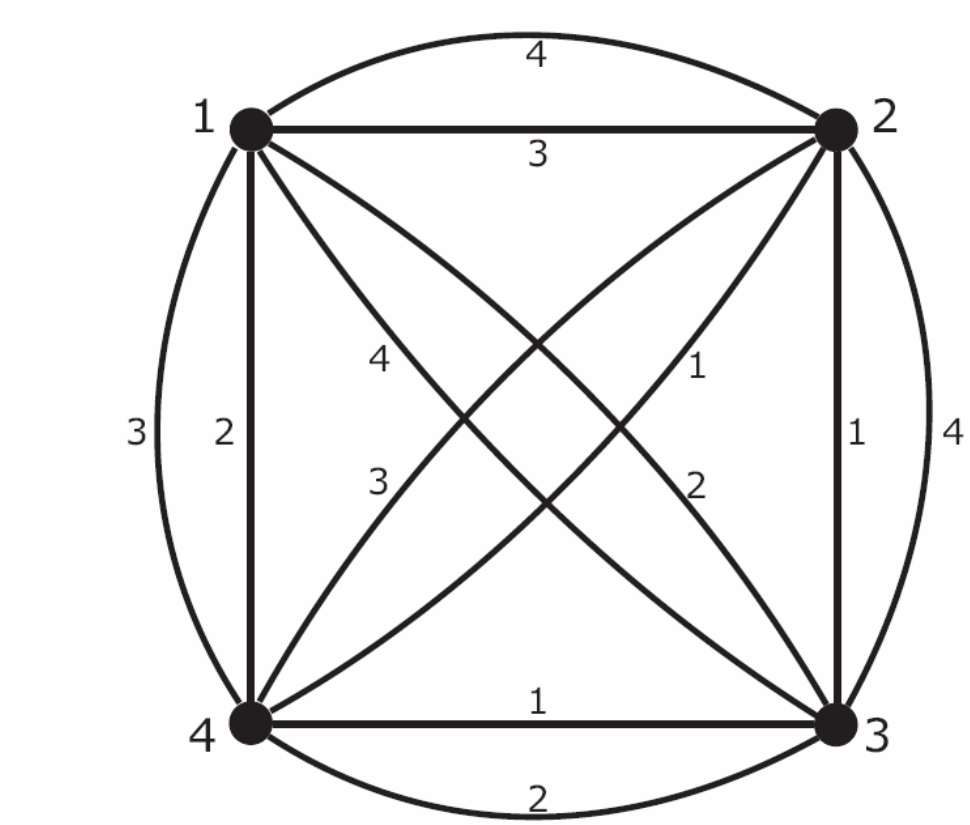


An example graph representing a system with four units

MM Comparison Model

- The edges indicate comparisons between a specific pair of units

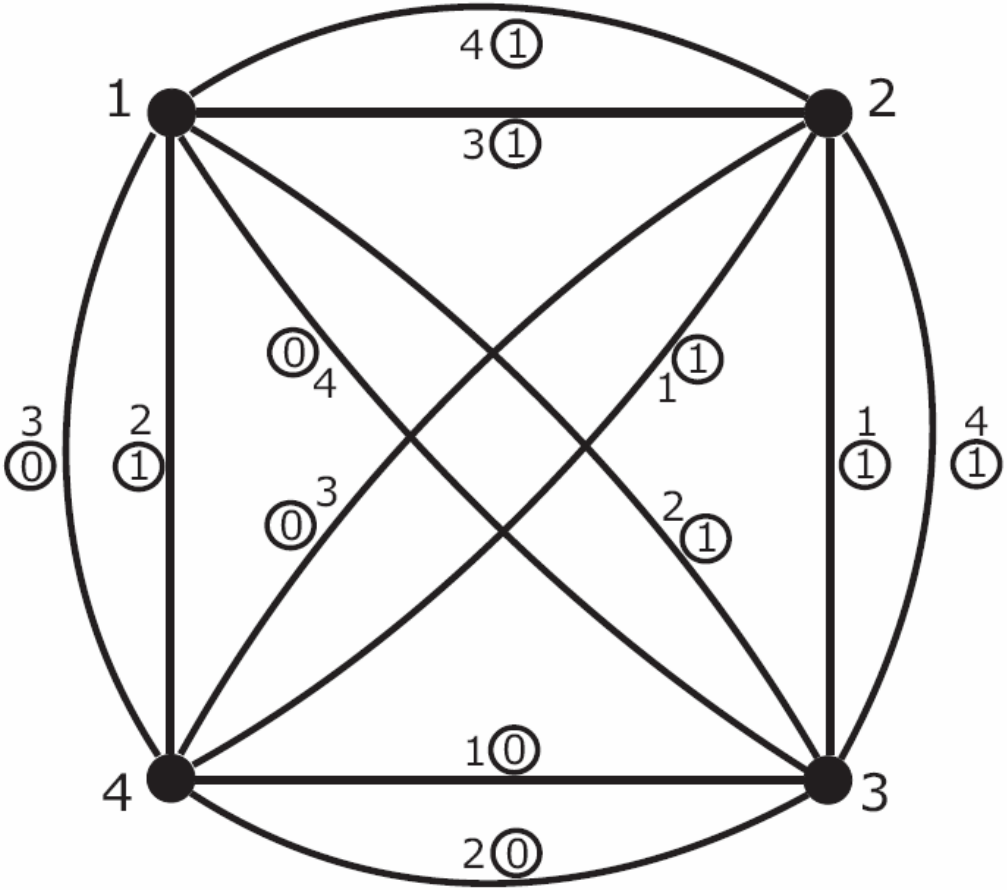
- Edge labels are id‘s of comparator units



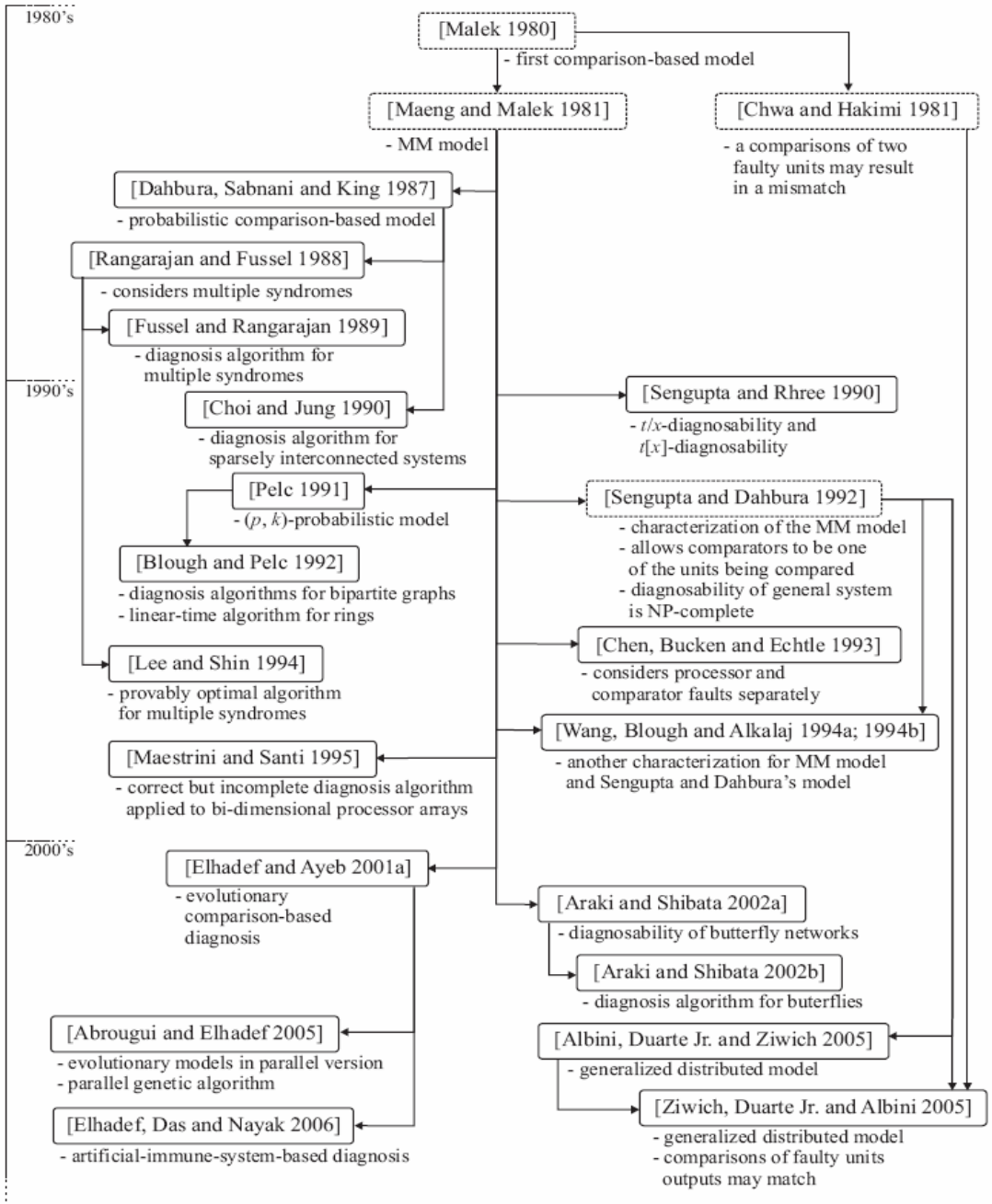
An MM-comparison multi-graph *M* for a system with four units

MM-Comparison Model

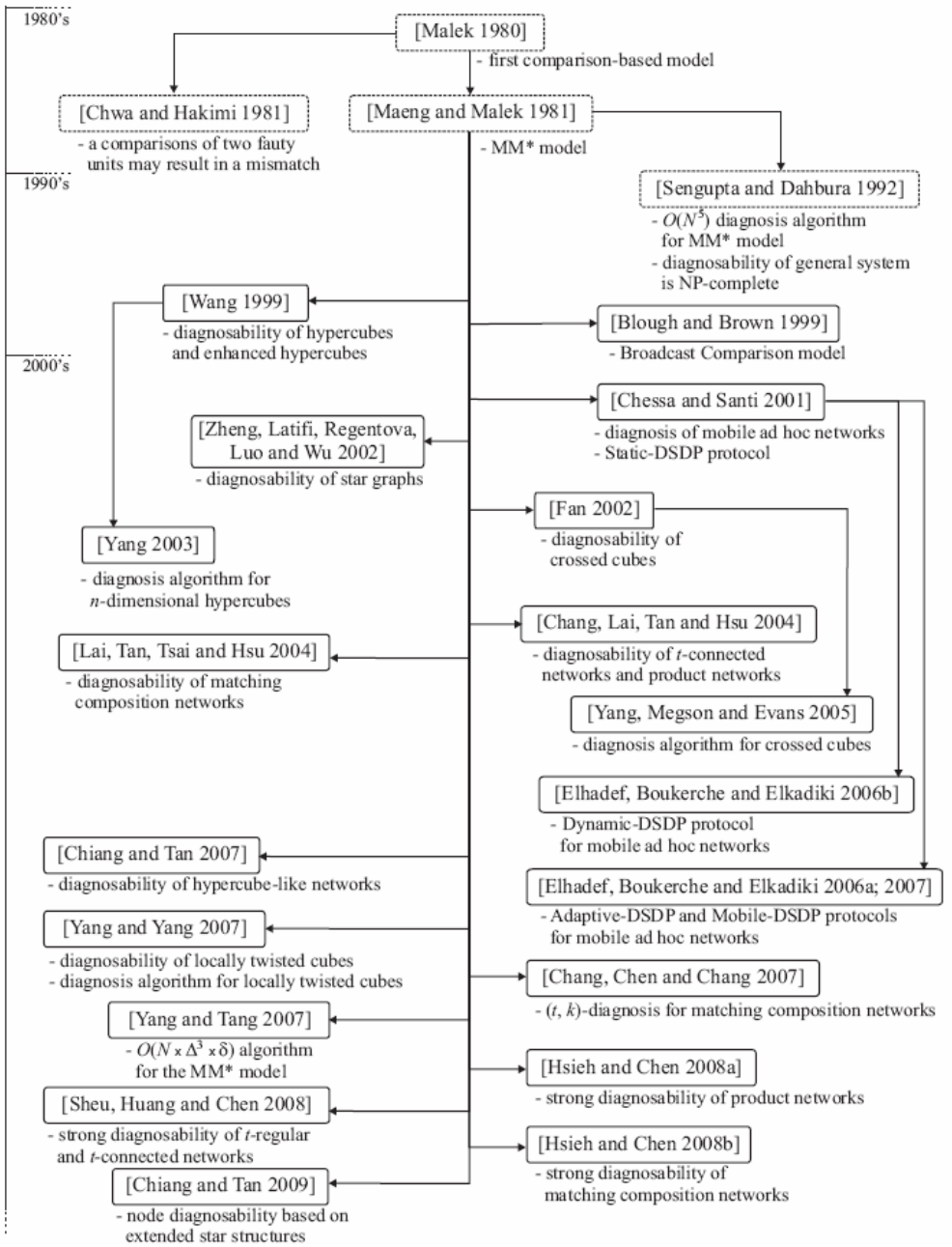
- Necessary and sufficient conditions for one-step diagnosability are given
- Algorithm for design of to diagnosable systems has been proposed
- Polynomial diagnosis algorithms (e.g., Sengupta and Dahbura)



Multi-graph *M* depicts comparison outcomes for the example system



Comparison-based diagnosis timeline: results based on the MM model from Duarte, Roverli, Ziwich, Albini, 2010



Comparison-based diagnosis timeline: results based on the MM* model from Duarte, Roverli, Ziwich, Albini, 2010

Main Directions

- Variations on assumptions
- Diagnosis algorithms
- t-diagnosability

Variations on assumptions

- Two faulty units may give identical outputs (Chwa and Hakimi)
- Probabilistic diagnosis (Masson, Dahbura et al)
- Distributed diagnosis (Kuhl and Reddy, …)
- Reliability of communication
- Reliability of comparators/processors
- …

Bottom line:

Let's get a consensus on minimal and realistic assumptions

Diagnosis algorithms

- One-step and sequential diagnosis algorithms, centralized versus distributed
- O (n^2) distributed algorithm for the basic model (Amman and Dal Cin)
- O (n^5) diagnosis algorithm for the MM* model plus NP-completeness result (Sengupta and Dahbura)
- O ($n\ d^3_{max}\ d_{min}$) diagnosis algorithm (Yang and Tang)
- A number of other algorithms for specific topologies and applications (wireless)

t-diagnosability

- Several special cases with respect to specific topologies (mesh, hypercube, twisted cube, butterfly, etc)
- Several result regarding varaitions on t-diagnosability
  - t/(N-1)-diagnosability (Xu,Huang, Randell)
  - t/m-diagnosability (m misleading comparisons, Krawczyk)
  - t/x- and t/[x]-diagnosability (x missing c., Sengupta)
  - t/s and t_1 /t_1 -diagnosability (up to s= t_1 can be replaced, Friedman, Masson et al)

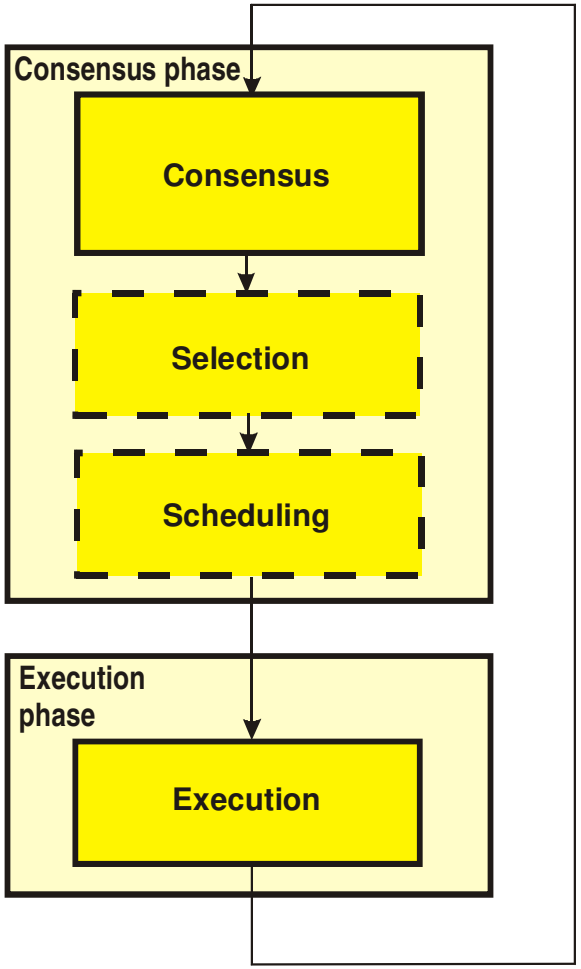
for either one-step or sequential diagnosability

Implementations

- Large number of applications

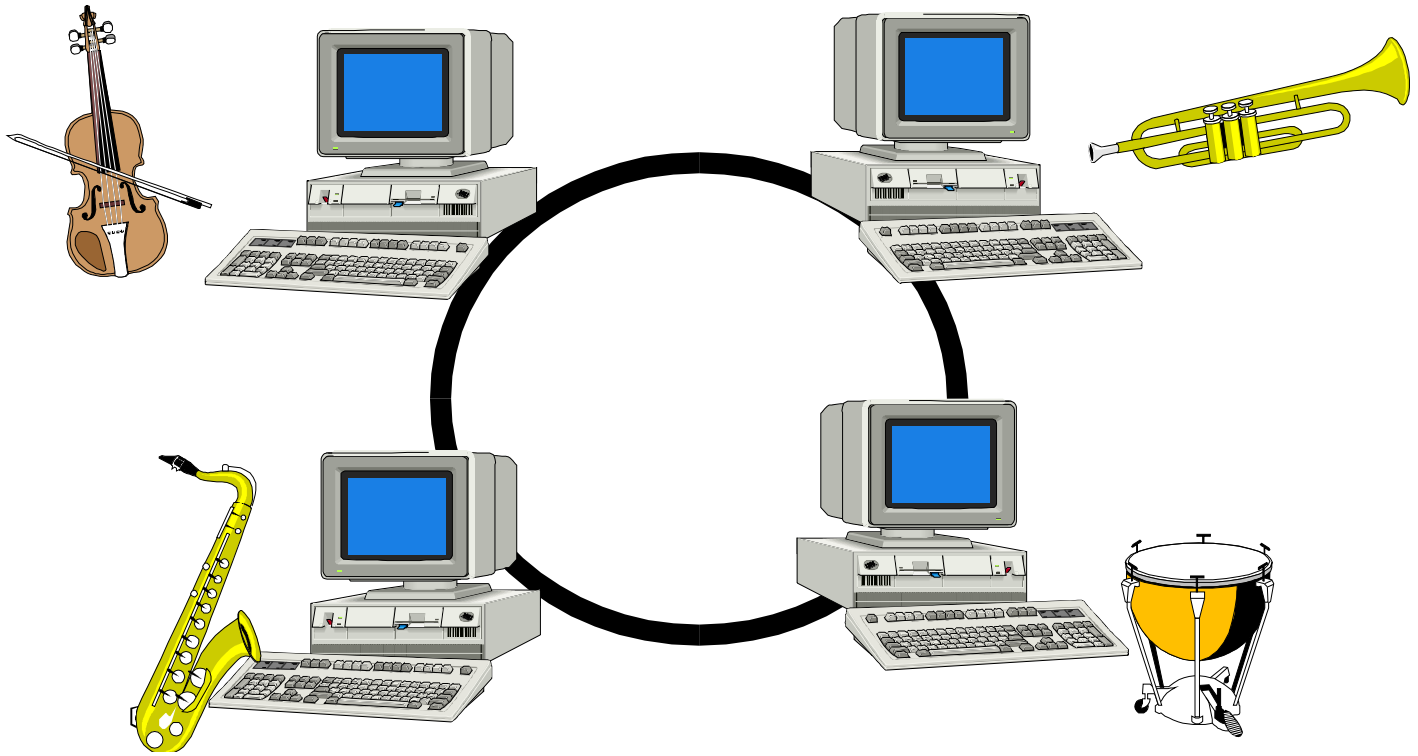
- This summary is from a personal perspective

CORE - COnsensus for REsponsiveness

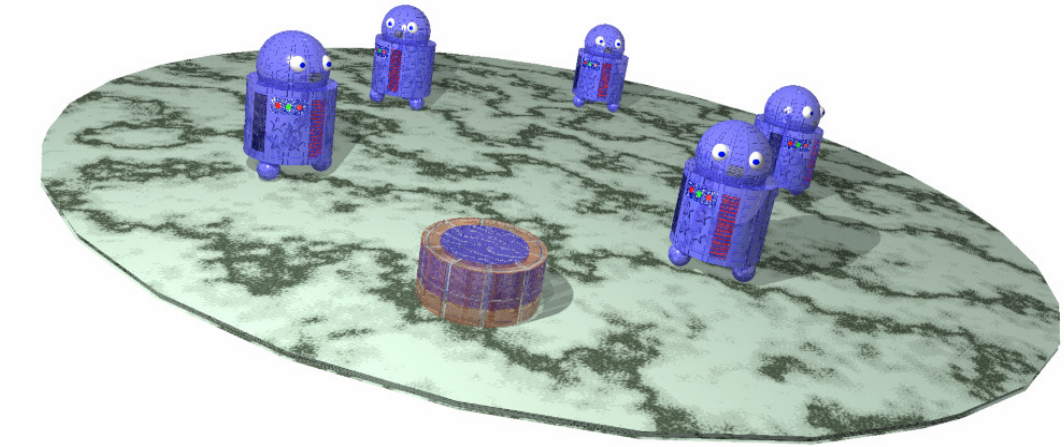


- Dependable architecture for distributed systems
- Alternating consensus/diagnosis phase and execution phase
- No communication during execution phase

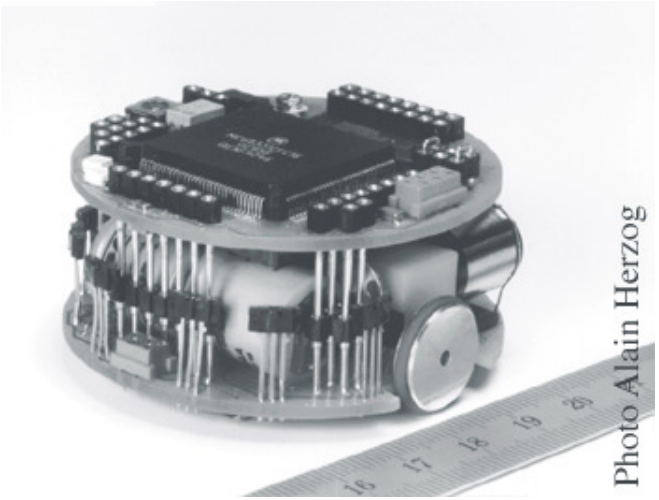
The Unstoppable Orchestra



Balancing the Robots

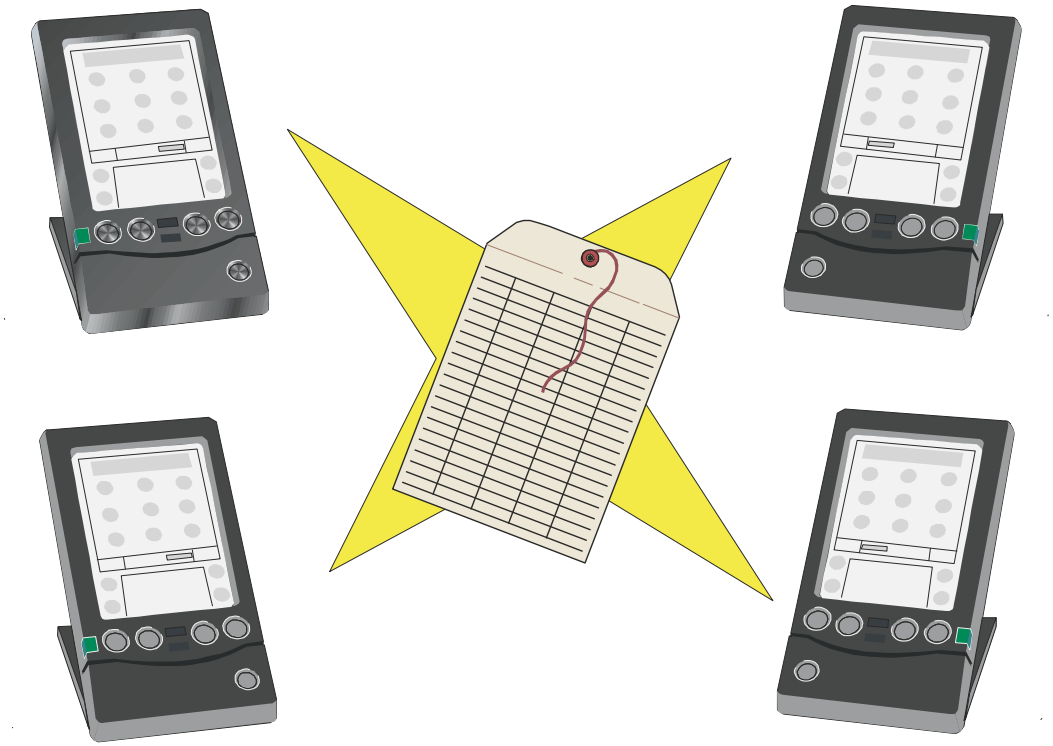


- Keeping an instable plate in balance
- A fault may immobilize a robot



Security by Consensus

- "Treasure-box" approach, agreement by comparison
- Data are accessible when a weighted majority agrees



Implementations

- Multiprocessor diagnosis at JPL (Wang, Blough, Alkalaj, 1994)

- Testing and diagnosis by comparison on the wafer without a golden unit (Rangarajan, Fussell and Malek, 1990, Agrawal, LogiTech)

- Mobile ad-hoc networks (Chessa and Santi, 2001, Elhadef, 2007

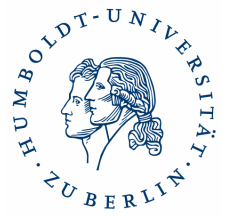
- Data integrity (Ziwich, Duarte and Albini 2005)

- Application-level diagnosis for generic time-triggered systems (Serafini et al 2010, Suri, Kopetz)

The System Diagnosis Questions

- Fault models (active nodes only, active and passive nodes, synchronization, frequency)
- Centralized or distributed, hierarchical
- Detection, location, fail-over, recovery
- Coverage, granularity, level, scalability and speed
- Static versus dynamic methods

Challenges in the Context of Comparison-based Methods



- What are the most realistic models and assumptions?
- What features/variables should be compared to make diagnosis most effective at each level?
- How to minimize monitoring and comparison overheads, synchronization and frequency?
- Dealing with diversity of HW, SW, people, etc.
- Dealing with uncertainty of comparisons?
- Diagnosis of temporary faults and new problems
- Can exotic faults such as configuration faults be handled by comparison? Encoding configuration
- Keeping it simple

Future Applications

- Cloud and grid computing
- Multicore and many core systems
- Comparison in new communication environments (especially wireless)
- Data integrity and security
- Embedded systems, sensor networks

Appendix:
A summary of MM and MM* models

Model	Reference	Main Contributions
MM model	[Maeng and Malek 1981]	- comparison diagnosis model in which units are also comparators - comparison outputs when at least one unit is faulty always results in a mismatch - central observer is a trusted unit that performs diagnosis - necessary and sufficient conditions for one-step t-diagnosability - procedure to construct minimal graph for diagnosable systems - evaluation of diagnosis latency in terms of test cycles
	[Sengupta and Dahbura 1992]	- generalization of the MM model: allows comparators to be one of the units being compared - characterization of diagnosable systems under the MM model - diagnosability of general systems is NP-complete
	[Sengupta and Rhee 1990]	- t/x-diagnosability and $t[x]$-diagnosability
	[Chen et al. 1993]	- extension of MM model considering processor and comparator faults separately; diagnosability evaluation
	[Wang et al. 1994a; 1994b]	- new necessary and sufficient diagnosability conditions for both the MM model and Sengupta and Dahbura's model
	[Maestrini and Santi 1995]	- correct but incomplete diagnosis algorithm applied to locate faults in bi-dimensional processor arrays
	[Araki and Shibata 2002a]	- diagnosability of k-ary r-dimensional butterfly networks
	[Araki and Shibata 2002b]	- $O(k^2n)$ diagnosis algorithm for butterfly networks
MM* Model	[Maeng and Malek 1981]	- MM* model is a special case of the MM model: each unit compares all pairs of neighbors
	[Sengupta and Dahbura 1992]	- diagnosis algorithm with time complexity $O(N^5)$ under the MM* model - diagnosability of general systems under the MM* model is NP-complete
	[Yang and Tang 2007]	- diagnosis algorithm with time complexity $O(N \times \Delta^3 \times \delta)$ under the MM* model, where Δ and δ are respectively the maximum and the minimum degrees of a node
	[Wang 1999]	- diagnosability of hypercubes and enhanced hypercubes
	[Yang 2003]	- worst case $O(Nlog_2^2N)$ diagnosis algorithm for hypercubes
	[Fan 2002]	- diagnosability of crossed cubes
	[Yang et al. 2005]	- $O(Nlog_2^2N)$ diagnosis algorithm for crossed cubes

(from Duarte, Roverli, Ziwich, Albini, 2010)

Model	Reference	Main Contributions
MM* Model *(continued)*	[Yang and Yang 2007]	- diagnosability of locally twisted cubes - $O(Nlog_2^2N)$ diagnosis algorithm for locally twisted cubes
	[Chiang and Tan 2007]	- diagnosability of hypercube-like networks
	[Zheng et al. 2002]	- diagnosability of star graphs
	[Lai et al. 2004]	- diagnosability of matching composition networks
	[Chang et al. 2007]	- (t, k)-diagnosis for matching composition networks
	[Chang et al. 2004]	- diagnosability of t-connected networks - diagnosability of product networks
	[Sheu et al. 2008]	- strong diagnosability of t-regular and t-connected networks
	[Hsieh and Chen 2008a]	- strong diagnosability of product networks: hypercubes, mesh-connected k-ary n-cubes, torus-connected k-ary n-cubes, hyper-Petersen networks
	[Hsieh and Chen 2008b]	- strong diagnosability of matching composition networks: n-dimensional crossed cubes, Möbius cubes, twisted cubes and locally twisted cubes
	[Chessa and Santi 2001]	- comparison-based diagnosis applied for mobile ad hoc networks - Static-DSDP protocol for fixed topology
	[Elhadef et al. 2006b]	- protocol Dynamic-DSDP for ad hoc networks based on Chessa and Santi's model
	[Elhadef et al. 2006a; 2007]	- comparison-based diagnosis applied for mobile ad hoc networks - Adaptive-DSDP Protocol for fixed topology networks - Mobile-DSDP protocol for time-varying topology networks
	[Chiang and Tan 2009]	- node diagnosability based on extended star structures
Broadcast Comparison Model	[Blough and Brown 1999]	- fully distributed comparison model - based on MM* for systems with reliable broadcast - polynomial-time algorithms to diagnose static and dynamic fault situations
Generalized Distributed models	[Albini et al. 2005; Albini and Duarte Jr. 2001]	- the generalized distributed comparison-based model: a hierarchical, adaptive and distributed model based on Sengupta and Dahbura's model - *Hi-Comp* diagnosis algorithm: requires at most $O(N^3)$ comparisons and has worst-case latency of $O(log_2N)$ rounds
	[Ziwich et al. 2005]	- generalized distributed comparison-based model assuming the comparison of faulty units outputs may match - *Hi-Dif* diagnosis algorithm that requires at most $O(N^2)$ comparisons and has worst-case latency of $O(log_2N)$ latency

(from Duarte, Roverli, Ziwich, Albini, 2010, continued)

Model	Reference	Main Contributions
Probabilistic model	[Dahbura et al. 1987]	- probabilistic comparison based model - considers probabilities for a match or a mismatch when comparing units
	[Rangarajan and Fussell 1988]	- strategy based on the evaluation of multiple syndromes
	[Fussell and Rangarajan 1989]	- $O(log_2 N)$ for the evaluation of multiple syndromes
	[Lee and Shin 1994]	- probably optimal algorithm for the evaluation of multiple syndromes
	[Choi and Jung 1990]	- diagnosis algorithm for sparsely interconnected systems
(p, k)-Probabilistic model	[Pelc 1991]	- a task has k possible outcomes - each unit has the same probability $p < 1/2$ - probability of obtaining a match when comparing a faulty unit and a fault-free unit or two faulty units is $q = 1/k$ - *diagnosis* and the *diagnosability* problems are NP-hard for general topology
	[Blough and Pelc 1992]	- polynomial time diagnosis algorithms for bipartite graphs (includes hypercubes, grids and forests) - linear-time algorithm to perform optimal diagnosis of rings
Evolutionary Comparison-Based models	[Elhadef and Ayeb 2001a]	- evolutionary comparison-based diagnosis
	[Abrougui and Elhadef 2005]	- parallel evolutionary diagnosis models
	[Elhadef et al. 2006]	- comparison-based diagnosis model with an artificial-immune-system-based approach

(from Duarte, Roverli, Ziwich, Albini, 2010 , continued)