

Independent Administrative Institutions and Roles of Universities

-Past, Present, and Future-
(Revised Version)

Tosiyasu L. Kunii

Professor Emeritus of the University of Aizu
Professor Emeritus of the University of Tokyo
Life Fellow of IEEE

Independent administrative institute

- Many Japanese universities are struggling on survival games after public universities have been transformed to independent administrative institutes.
- Private universities have received cascading effects and also are in heavy survival games too. It is very unwise view to take a survival game view.

Placing universities on the right tracks

- A clear vision is needed to place universities on right tracks to let them play proper social roles both nationally. and internationally.
- This talk first clarifies the situation based on the historical information on the roles of universities at national and international levels, and then surveys current and future roles and potentials of universities.
- Some concrete and executable plans are presented based upon those at the University of Tokyo.

Major Global Social Human Resources

- Economic divide and IT divide having the major human resource on the uncared side of it indicate the best use of higher education.
- **One hundred dollar computer projects** such as **OLPC** (one laptop per child) project aim at the **numbers** for the uncared side.
- Their results depict now the **educational quality** satisfy naturally far higher demand of the uncared side.

The most sound businesses for IAI.

- Further, it produces the most valuable high level human resources to make it among the most sound businesses for IAI.
- The secure and valid business models are essential to succeed as seen in the global failure of the subprime model and the following off-balance model and SIV (Structured Investment Vehicle).

Independent Administrative Institutions (IAI)

- So, transforming “IAI” into the social and international global units to meet the demand raises the best outcome.
- It has to be **insured** with multiple feedback loops based on the cash flow distribution in time and space and expectations in short-, middle-, and long-terms. This is the basic of the business models for any business.

**The main framework:
Cyberworlds as the Leading
Power of World Evolution**

Cyberworlds

- ♣ Being formed as **information worlds** in cyberspaces that is a type of computational spaces, with or without design.
- ♣ Widespread and intensive **local activities** are melting each other on the web **globally** to create cyberworlds.

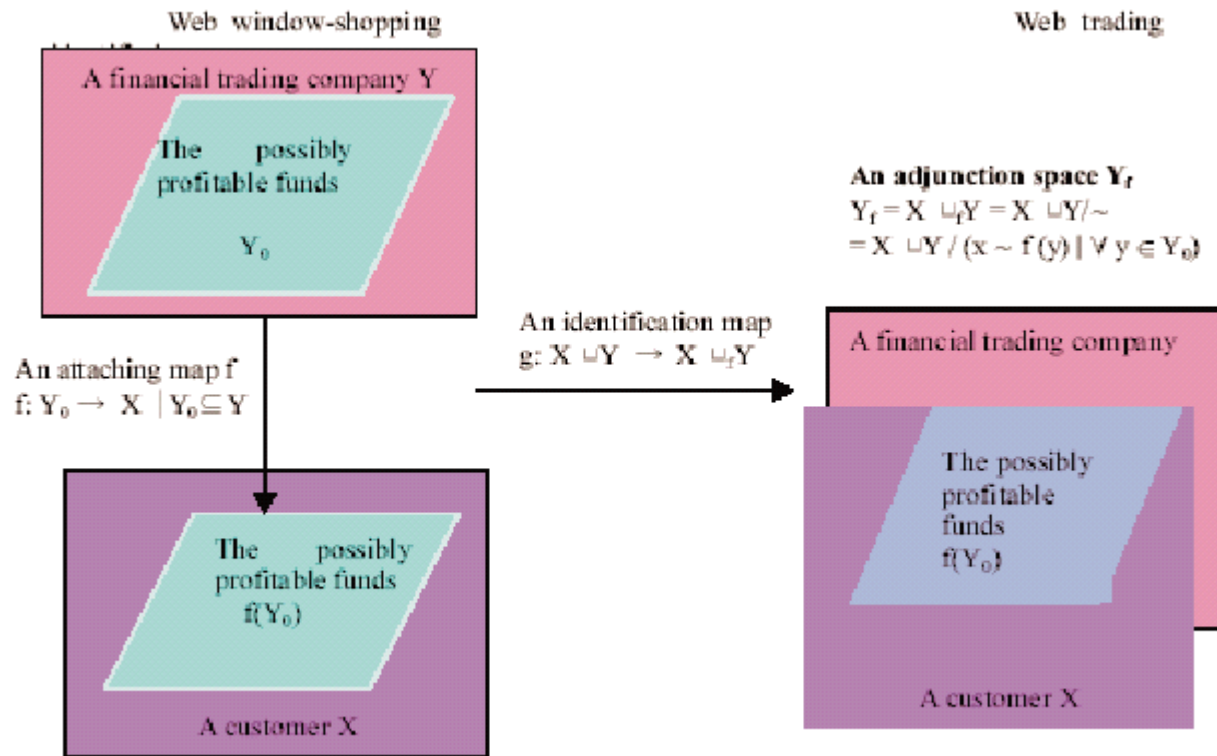
My discovery of cyberworlds

- ***In 1969, discovering cyberworlds in computational space while investigating the properties of computationally synthesized non-existent molecular worlds to find the molecular origin of life.***
 - ***Tosiyasu L. Kunii, “Invitation to System Sciences - Poetry, Philosophy and Science in Computer Age—”, (in Japanese), Journal of Mathematical Sciences, pp. 54-56 (October 1969), Science Publishing Co. Ltd., Tokyo, Japan.***
- ***Computer Science Department has been created at the University of Tokyo to study its nature in 1970 (originally as a research institute).***

The major key players of cyberworlds

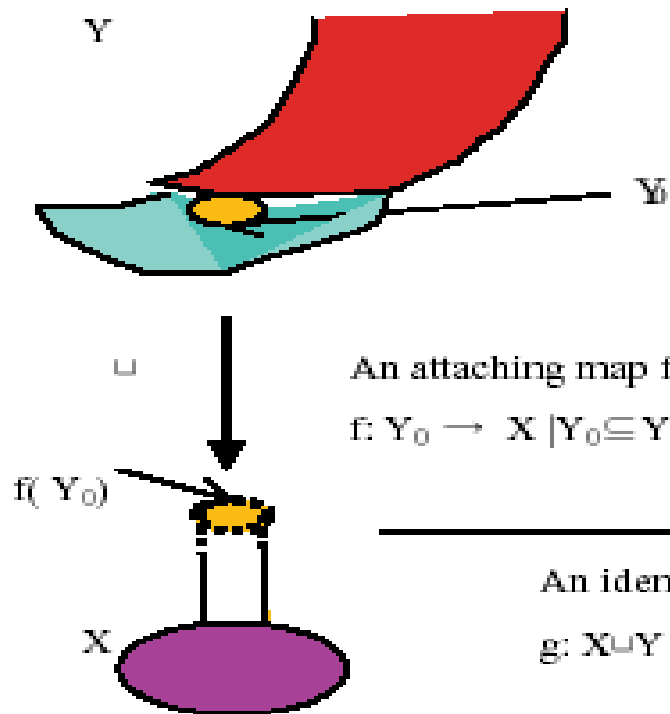
- ♣ **e-finance** that trades a GDP-equivalent a day.
- ♣ **e-manufacturing** that is transforming industrial production into Web shopping of
 - ♣ product components and
 - ♣ assembly factories.

Financial trading processes on the Web



e.g. e-manufacturing: a simple example

Disjoint topological spaces $X \sqcup Y$

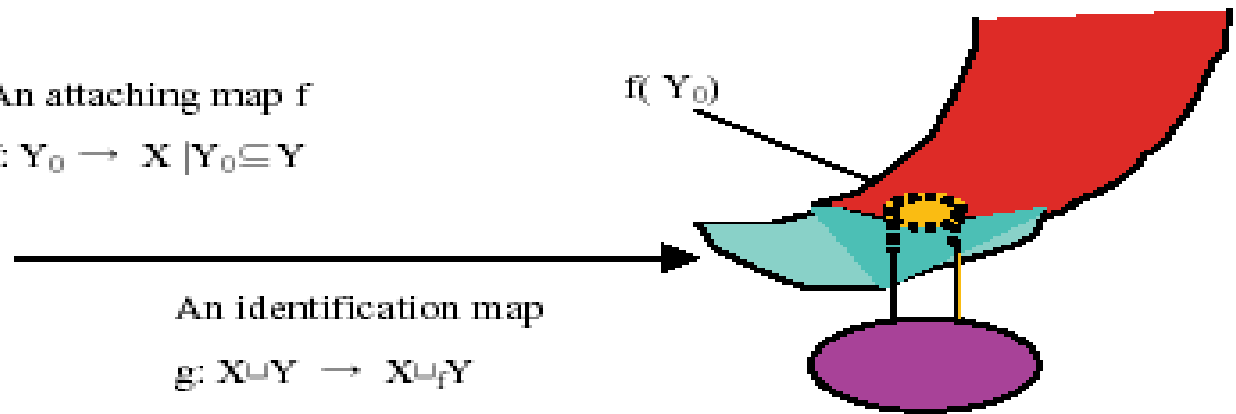


An attaching map f
 $f: Y_0 \rightarrow X \mid Y_0 \subseteq Y$

An adjunction space Y_f

$$Y_f = X \cup_f Y = X \cup Y / \sim$$

$$= X \cup Y / (x \sim f(y) \mid \forall y \in Y_0)$$



An identification map
 $g: X \cup Y \rightarrow X \cup_f Y$

Without proper modeling ...

♣ cyberworlds have continued to grow **chaotic** and are now out of human understanding and control.

What commonly required in macroscopic and microscopic cases to realize such potentials

- a provision to identify objects in cyberworlds and human cognitive spaces.

Modelling cyberworlds and cognitive spaces by an incrementally modular abstraction hierarchy

- *Identification* by continuous and surjective mapping named *identification mapping* (*quotient mapping*) in adjunction spaces.

A remedy: An adjunction space model

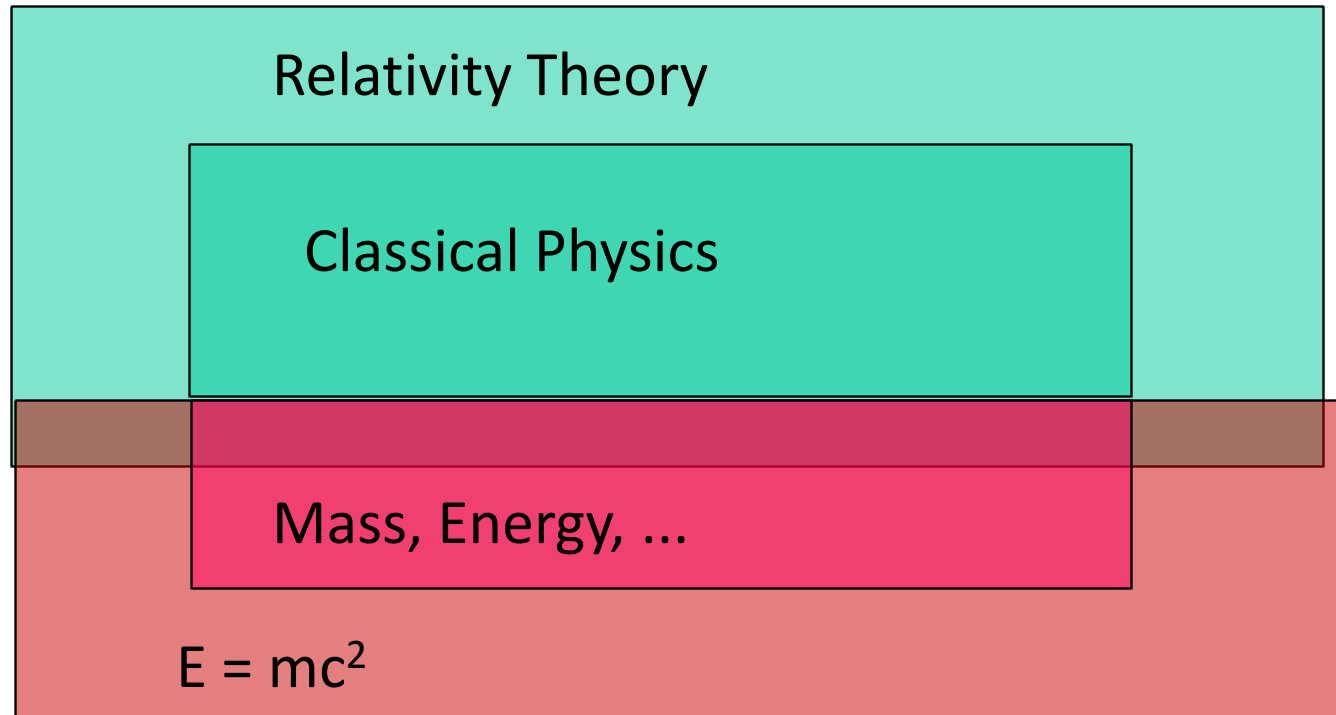
- ♣ To globally integrate local models;
- ♣ To capture spatio-temporal aspects of irregular information worlds as an information model;
- ♣ Mathematically, it is based on an incrementally modular abstraction hierarchy including cellular spatial structures in a homotopy theoretical framework.

What to identify ?

- Invariants by cognizing common unchanging properties.
- Mathematically, 'common' means **equivalent**, and what are common form an **equivalence class**.
- There are hierarchies of invariants.
- They stand for abstraction hierarchies.

A hierarchy of invariants: a physics case

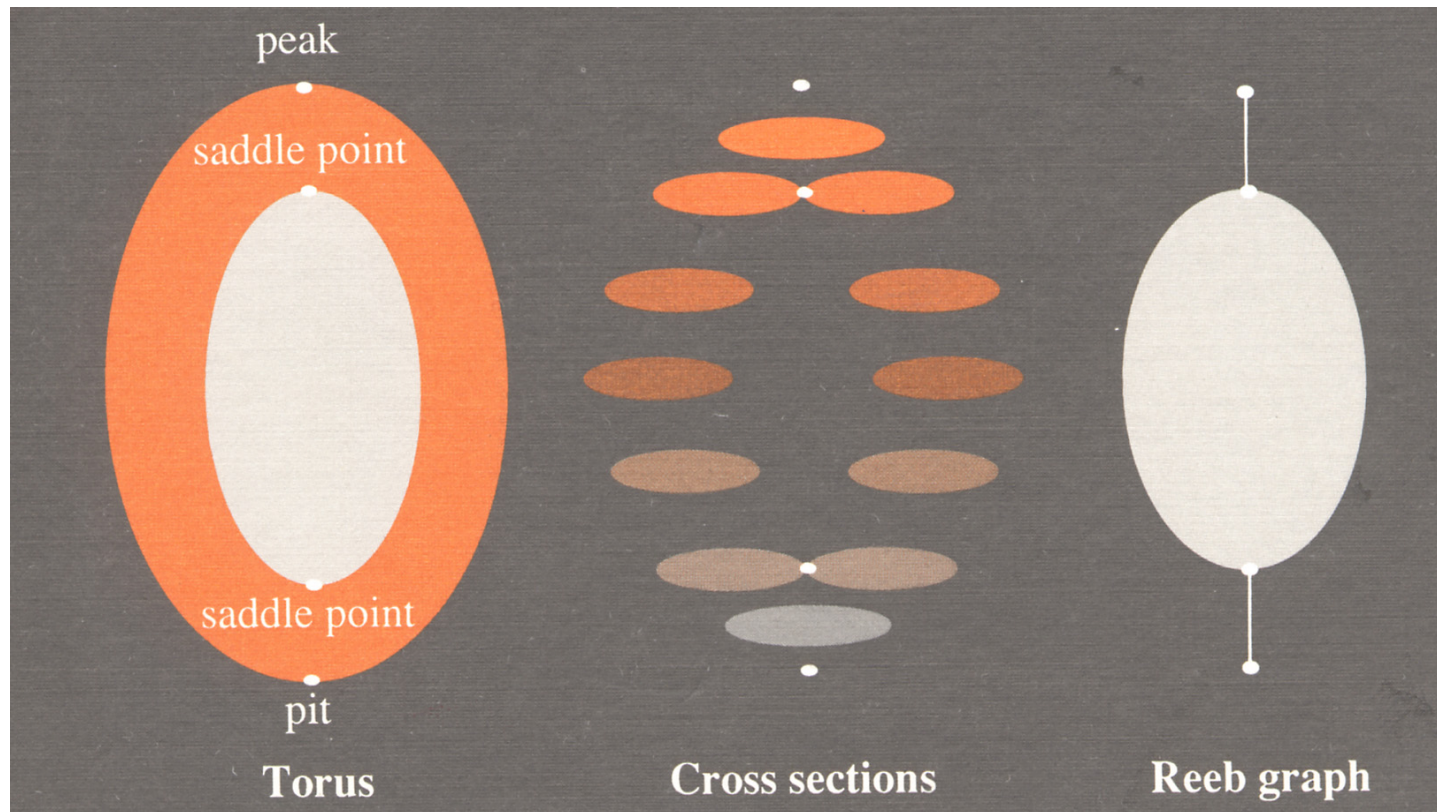
- Mass, Energy, ... $\implies E = mc^2$



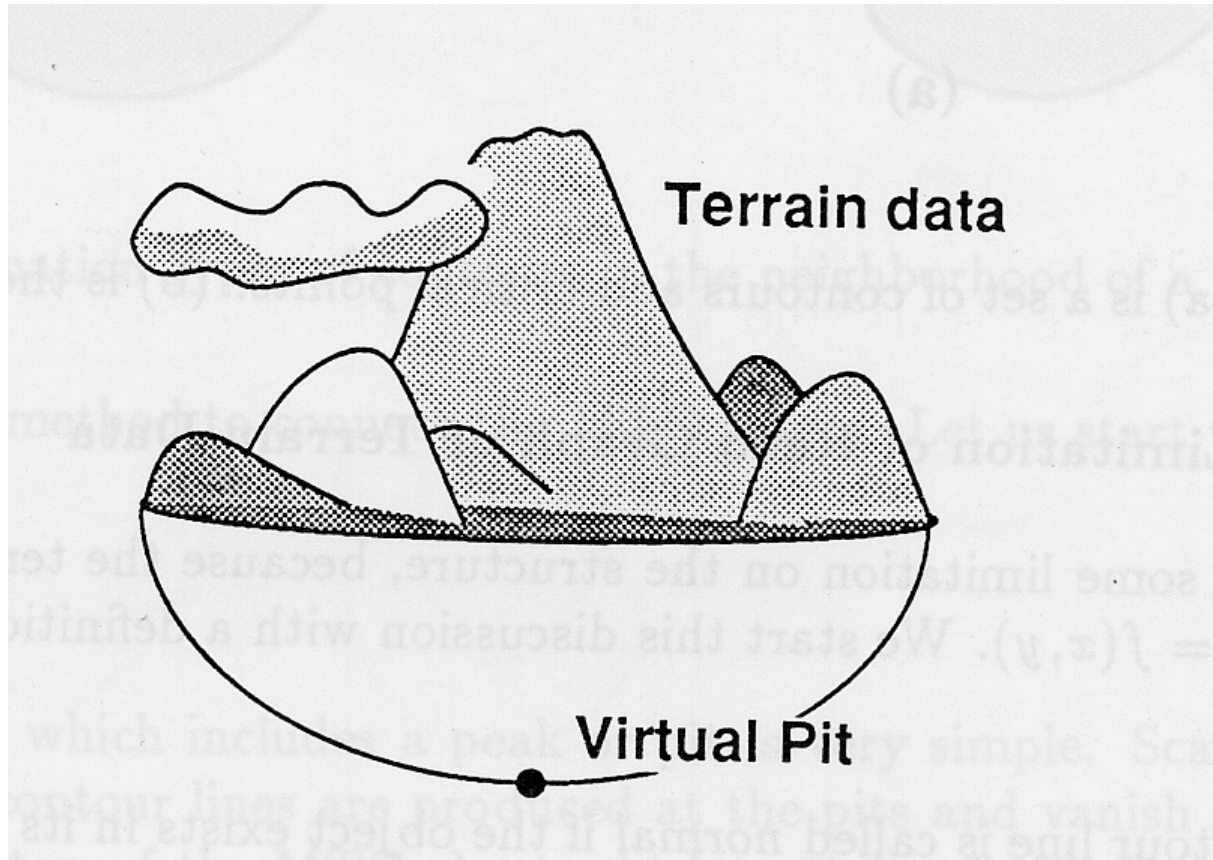
Invariants in Visualization

- Morse theoretical critical points such peaks, pits passes serve as invariants in differential topology.

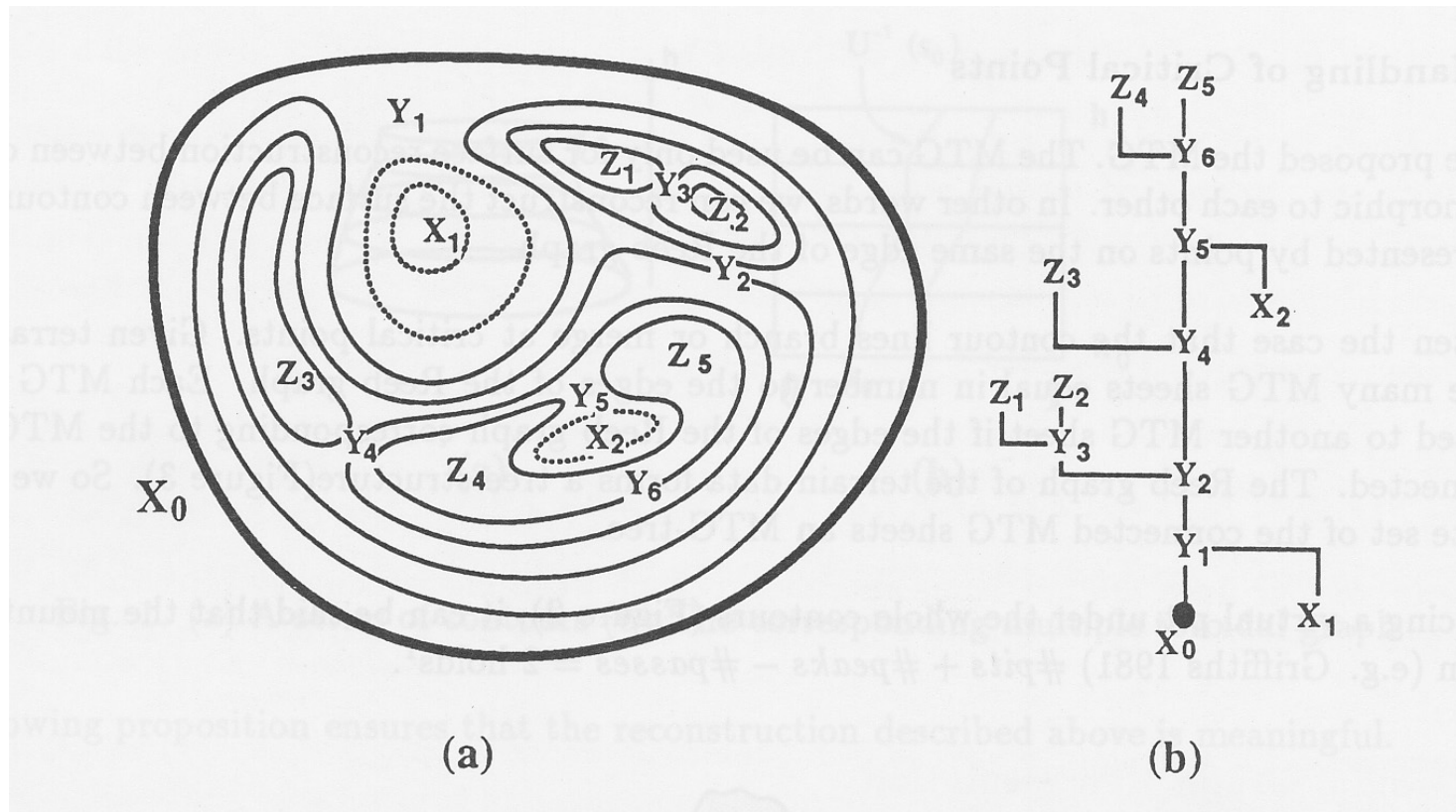
Differential topological design: Morse theoretical model and Reeb graph model



A Terrain

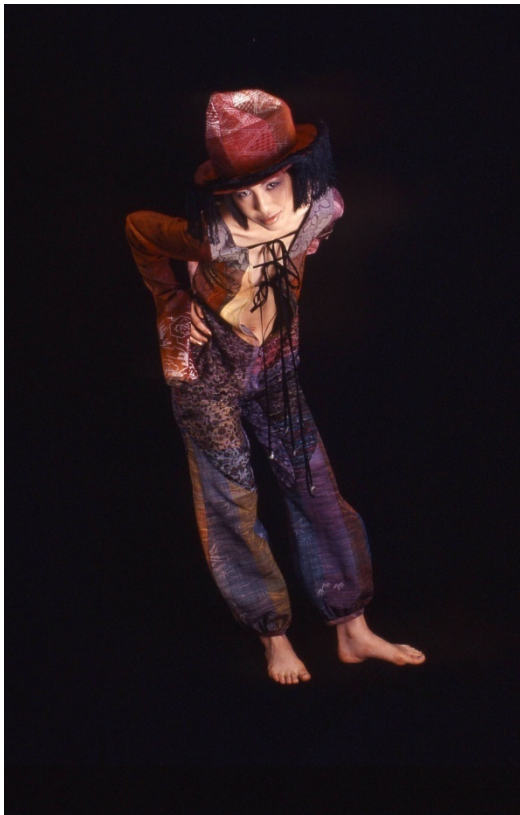


An Equihight Contour Lines, Critical Points, and the Reeb Graph



Automated in-between generation from a pair of pictures

Give a pair of pictures:



Considering the grey scale as the height function

- 1. The critical points as invariants are computed for each frame of pictures.
- 2. By matching the invariants of two frames of pictures, in-betweens are derived automatically to animate the picture.

To animate the picture



Let us turn growing cyberworlds into major social assets by making them reusable

♣ How?

- ♣ Via invariants preservation as seen in the example above.
- ♣ Via an incrementally modular abstraction hierarchy.

incrementally modular abstraction hierarchy

- ♣1. homotopy level;
- ♣2. The set theoretical level;
- ♣3. The topological space level;
- ♣4. The adjunction space level;
- ♣5. The cellular space level;
- ♣6. The representation level;
- ♣7. The view level.

We explain the hierarchy later.

How the incrementally modular abstraction hierarchy works?

- 1. To decrease the combinatorial complexity of cyberworlds design.*
- 2. The top-down design based on the hierarchy eliminates the combinations.*
 - to realize linear integration and linear interoperability.*

Impact projects at the time of world-wide recession

- “Web information management systems are emerging as key players in the global society we live (Tosiyasu L. Kunii and Annie Luciani, Eds., “Cyberworlds”, Springer-Verlag, 1998.).
- Continued gulf wars including the Iraq war, ever prevailing international terrorism and SARS have been spreading economic havoc, forcing companies and governmental organizations to merge for survival as mega corporations and digital governments.

Organization-wise merges

- Organization-wise merges do not save anything much any longer. It rather simply aims at eliminating further workplaces, resulting in more and more lay-offs to make the economy worse.
- First of all it does *not* usually mean organizational **operation** merge. The main reason is that the current information system technology provides element-by-element operation merge with inevitable **combinatorial explosion**.

The project that has the largest impact factor

- It is well known that the **combinatorial explosion** is the source of the sky high cost coupled with almost infinite delay of the operation merge after the organizational merge.
- Thus, *since the expectation of prolonged recession causes more and more organizations to merge*, the project that has the largest impact factor is that on **non-combinatorial organizational operation merge**. As such, the ideal is **linear**.

The combinatorial explosion of organizational operation merge: How it's brought in?

- The result of the individual designers working hard to manually create their “best” interfaces for the merge.
- Such “hard work” creates **different designs** for “best” interfacing of the different information systems in charge of individual organizational operations.

Similarity with the pre-data independence

- The situation is very similar to that when the relational model was proposed by E. F. Codd in 1970.
- At that time, commercial database management systems (DBMS) were pointer-based such as CODASYL and IBM IMS.
- When data changes occurred at some organization, the databases had to be updated manually.
- At individual customer sites, employees of DBMS companies were sort of sold to the customer sites spending the whole lives for manual data- and pointer-updates.
- They did their “best” to update the pointer design manually.

Reactions to the data independence

- That tragedy was saved by data independence of the relational model.
- The trend to refuse to recognize the value of the relational model in data independence had continued for at least a decade.
- I myself in supporting data independence had been treated as a **foe of the DBMS community** by experts, although I have organized Database Technical Committee and Research Group at Information Processing Society of Japan in 1972 and served as Vice-Chairman, and was a member of IFIP Working Group 2.6 on Data Bases soon after (1974 to 1992).

Lack of modeling for information system integration

- For information system integration, there is no model currently. Hence, the combinatorial explosions.
- One of the cases causing the combinatorial explosions is what's called “wrappers”. Wrappers can wrap anything in any style, creating varieties of “flexible” designs and ending up with remarkable design combinations to explode.

The way I was on information system integration since 1974

- 1. A **cumulative data model**: An inductive scheme (1974).
- 2. Application-, data-, and machine-independence model (1977).

A Theoretical Origin

- Theoretically identical with visual information systems
 - in that visual information needs to be identified and inter-related to turn necessary parts into reusable resources *without any initial world model* as clarified in 1974 as a *cumulative data model*. \implies 1974 discussion with Dr. E. F. Codd
 - see “A Relational Data Base Schema for Describing Complex Picture with Color and Texture”, T. L. Kunii, S. Weyl and J. M. Tenenbaum, Proc. of the 2nd International Joint Conference on Pattern Recognition, pp.310-316 (Lyngby-Copenhagen, August 1974) [also available as Stanford Research Institute Technical Note 93, SRI Project 8721 (June 1974); reprinted in Policy Analysis and Information Systems, Vol. 1, No. 2, pp. 127-142 (January 1978).

The Codd's Initial World Model

- The functional dependency of the Codd' relational model breaks down by getting a new instance, as shown in 1974 (“A Relational Data Base Schema for Describing Complex Picture with Color and Texture”) as a *cumulative data model*.
- Dr. E. F. Codd clearly replied that all the functional dependencies are assumed to be predefined in the relational model as the *initial world model* and hence a *cumulative data model* is outside the relational model.

Automatic integration of multiple information systems: up to the point in time of integration

- ♠ To avoid laborious manual work of renormalization, I have devised an **automated renormalization scheme** based on mechanical inductive inference on cumulative image database up to the point in time [Kunii74].
- ♠ It is equivalent to joining all the relations to turn them into one relation and mechanically induce functional dependencies for automated renormalization.
- ♠ The same situation arises when we integrate multiple information systems where up to the point in time of integration, we can automate renormalization of integrated database by the same inductive scheme.
- ♠ \implies Very time consuming yet!

Another Theoretical Origin

- Interconnecting heterogeneous databases of distributed computers: data conversion vs. query conversion
 - "Design Criteria for Distributed Database Systems", T. L. Kunii and H. S. Kunii, Proc. of the third International Conference on Very Large Data Bases (VLDB 77), pp. 93-104(Tokyo, Japan, October 1977).

UBITA: The minimum Upper Bound of Information Translation Amount

in database sharing as a criterion for the compositional design approach.

- Different design models are tested against these design criteria, by taking typical distributed database system environments. In pursuit of highly adaptive design for very general evolving heterogeneous distributed database systems, an **abstract design approach** is proposed.
- It is expected that this design also minimizes the system maintenance load which has already taken up with much more man power than the system development work.

Application-, Data-, and Machine-Independence criteria

- In this approach, Applications (A), Data (D), Machines (M), and a system Controller (C) are identified as major logical subsystems of an evolving heterogeneous distributed database system. Two sets of three independence's (3Is, that is, Application-, Data-, and Machine- Independence's) criteria are proposed for adaptive evolving heterogeneous system design:
 - (i) 3Is for a non-cascading system design to prevent a change in any one of the A-, D-, and M- subsystems from cascading into the other two logical subsystems, and
 - (ii) 3Is for an invariant system controller to keep the system controller invariant under any *change* in the A-, D-, or M- subsystems.
- The *change* includes **information system integration**.

A virtualization method

- A virtualization method is discussed as a way to fulfill the two sets of the 3Is criteria.
- It is also shown that the conventional data independence is insufficient for adaptive system design including integrating information systems.

Meet the application-, data-, and machine- independence criteria

- It is extremely hard to meet the application-, data-, and machine- independence criteria.
- Theorem: Adjunction spaces represent the information common to all applications, data and machines to meet the application-, data-, and machine- independence criteria.

An intuitive proof

- The only way to achieve the goal is to find a very high level of abstraction mechanism to *abstract the information common* to all applications, data and machines.
- It has been 26 years to find the answer to meet the criteria by finding the meaning of “*common*” and what’s common as *adjunction spaces* in algebraic topology
[Kunii2003: Toshiyasu L. Kunii, Masumi Ibusuki, Galina I. Pasko, Alexander A. Pasko, Daisuke Terasaki, and Hiroshi Hanaizumi, “Modeling of Conceptual Multiresolution Analysis by an Incrementally Modular Abstraction Hierarchy”, Transactions of Institute of Electronics, Information and Communication Engineers, pp. 1181-1190, July 2003.].
- “To be *common*” means “to be *equivalent so that the equivalent information is shared to be common*”. Equivalence relations mathematically define the meaning of to be equivalent. Equivalent information forms adjunction spaces via identification functions on quotient spaces that are also called identification spaces. Then, adjunction spaces represent the information common to all applications, data and machines to meet the application-, data-, and machine- independence criteria. Q.E.D

Adjunction spaces

- It thus serves as a novel data model that can integrate information systems *linearly* and hence avoiding the combinatorial explosion of the integration workload.
- The adjunction spaces model the common properties of dominant commercial information systems being used by major private and public organizations by abstracting the *common properties to be equivalent among different information systems* as adjunction spaces.

Automated linear interface generation

- For automated linear interface generation after the linear integration at the adjunction space level, we use an incrementally modular abstraction hierarchy.

(Tosiyasu L. Kunii, “Web Information Modeling: The Adjunction Space Model”, Proceedings of the 2nd International Workshop on Databases in Networked Information Systems (DNIS 2002), pp. 58-63, The University of Aizu, Japan, December 16-18, 2002, Lecture Notes in Computer Science, Subhash Bhalla, Ed., Springer-Verlag, December, 2002 [Kunii2002a].)

An incrementally modular abstraction hierarchy

- An incrementally modular abstraction hierarchy [Kunii2002a] is as shown below such that we are interfaced to existing information systems to the extent we realize linear interoperability to perform the integrated system-wide tasks.

The adjunction space level of an incrementally modular abstraction hierarchy

- At the adjunction space level, existing information systems are interfaced to the extent we realize *linear interoperability* to perform the integrated system-wide tasks.
- 1. The homotopy level;
- 2. The set theoretical level;
- 3. The topological space level;
- 4. The adjunction space level;
- 5. The cellular space level;
- 6. The representation level;
- 7. The view level.

Implementation

- Rapid prototyping is being tested in Python.

The major problems

- How to work with dominant existing systems that have no clean interoperability provisions.
- The relational model is a typical example.

The relational model-based information systems

- Still they are the easiest to achieve the automated linear interface generation to integrate them at the adjunction space level,
- by extracting the *cellular space level* information of different information systems to achieve the linear interoperability.

ERP

- Enterprise Resource Planning (ERP) is also fairly easy to achieve linear automated integration by our approach because in formalism ERP is graph theoretical and easily transformed to process graphs supported by the relational model.

(Tosiyasu L. Kunii, “The 21st Century Manufacturing Automation and the Social Impact”, Proceedings of International Conference on Manufacturing Automation (ICMA’97), April 28-30, 1997, Hong Kong, pp.21-29 , A. T .Tan, T. N. Wong and I. Gibson, Editors (Department of Mechanical Engineering, University of Hong Kong)).

Comparison with the relational model, the entity relationship model, UML, and XML

- None of the currently popular information models such as the relational model, UML, and XML have the **information modularity** simply because they are not based on the disjoint unions of information.
- Hence, modular Web information modeling based on equivalence to identify necessary information does *not* apply to any of them.

A JV example

- Suppose on the Web a company X is searching for a joint venture partner company, say Y . After the successful search, a joint venture (JV) is formed such that $X \sqcup Y / (x \sim f(y) \mid \forall y \in Y^0)$ via an attaching function f where $f: Y^0 \rightarrow X \mid Y^0 \subseteq Y$. $Y^0 \subseteq Y$ is the part of company Y running the JV with the corresponding part $f(Y^0)$ of company X .
- As a model, this JV model belongs to, and is exactly identical with, the adjunction space model explained so far.

M&A and JV

- Now, let us suppose a company X is searching for a company, say Y , to merge and acquire (M&A) Y . After successful M&A, denoting the part Y^0 of Y ($Y^0 \subseteq Y$) merged to X , we get the union of X and Y joined by the part $Y^0 \subseteq Y$ common to that of Y such that $X \sqcup Y / (x \sim f(y) | \forall y \in Y^0)$.
- In the relational model, the join operation performs M&A, but **the relational model fails to support JV.**

ER, UML and XML

- The entity relationship (ER) model is basically an intuitive graph theoretical model, and the model itself fails to support the identification by equivalence. So is UML.
- XML is flexible, born from bibliographic markup language SGML. It consists of nested pairs of brackets, lacking formalisms to validate the ever-expanding and complicated nesting. When the system being constructed becomes very large as usual in any practices, it falls apart by the lack of the formal and hence automated mechanism of validation. We have to formulate XML in the half Hausdorff spaces T_0 .

What has been popular in general such as “wrappers” and what’s wrong with them?

- Standardizing without theorizing
- Modeling and designing as the first phase
- Wrappers

Standardizing

- Standardizing is on the assumption for the existing systems to be rebuilt to follow a given specification called the “standard”.
- For information systems as the infrastructure of major social organizations, the cost of the rebuilding is prohibitively high to make the assumption usually invalid.

Theorizing

- Posing theorizing as a counter-concept of standardization, let us examine it. If we define theorizing in the following way, it becomes valuable:
- Theorizing means reasoning on the required features of major information systems followed by *abstracting the common features* of such systems to find the consistent logical structures called theories.

Modeling and designing as the first phase

- Now, suppose we do not standardize. Then, if we make *the modeling and the designing the first phase without theorizing*, we end up with a combination of the models and designs of different information systems leading us to **combinatorial explosions**.

Wrappers

- Saying “interfaced by wrappers” does not mean more than “interfaced somehow”.
- Historically what meant by wrappers have been in more and more varieties including interfaces in general, and relational extensions as seen in [Stonebraker].
- (Michael Stonebraker and Paul Brown, “Object-Relational DBMSs – Tracing the Next Great Wave-“, Morgan Kaufmann Publishers, Inc., San Francisco, CA USA, 1999.)

Wrappers



Conclusions

- Researches are open ended by definition, hence have no conclusions. Let us continue on together.