Input-Output Types of Fifteen Modules on Discrete and Real Measurements for COVID-19

Jeffrey Zheng (✉ conjugatelogic@yahoo.com)

Key Laboratory of Quantum Information of Yunnan, Key Laboratory of Software Engineering of Yunnan, Yunnan University <https://orcid.org/0000-0003-4225-7077>

Minghan Zhu

Yunnan University

Research Article

Keywords: metagenomic analysis system, multiple modules, workflow, discrete measures, genomic index, geometric invariant

Posted Date: January 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-65158/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Input-Output Types of Fifteen Modules on Discrete and Real Measurements for COVID-19

Jeffrey Zheng, Minghan Zhu

Abstract Fifteen modules in the MAS provide unique functions to support wider applications. It is important to describe this set of transformation schemes from input genomes and output results using visualable numbers to briefly explain the transformation mechanism involved in each module. In this paper, vectors, matrices and various measurements are described following each module. From a genomic sequence, there are two types of outputs for three groups: 1) groups A & B related to discrete measurements represented in either vectors or matrices; 2) groups C related real measurements linked to entropy indices restricted in a limited line segment or a certain square region as a hierarchical organization. The main transformations and projections are described. Sample distributions and genomic index maps are illustrated.

Keywords: metagenomic analysis system, multiple modules, workflow, discrete measures, genomic index, geometric invariant

Jeffrey Zheng^{1,2}

¹ Key Laboratory of Quantum Information of Yunnan

² Key Laboratory of Software Engineering of Yunnan
Yunnan University, Kunming, e-mail: conjugatelogic@yahoo.com

Minghan Zhu

Yunnan University e-mail: crystalj000@163.com

This work was supported by the NSFC (62041213), the Key Project on Electric Information and Next Generation IT Technology of Yunnan (2018ZI002)

Introduction

It is essential for readers to understand the basic functions of transforming modules in the metagenomic analysis system (MAS). Fifteen modules of three projections on the $\{A,B,C\}$ groups in the MAS provide unique capacities to support wider applications. It is important to describe this set of transformation schemes from input genomes and output results relevant to vector or matrix using visual workflows to briefly explain the transformation mechanism involved in each module based on combinatorial algorithms, matrix theory, geometric measure theory, combinatorial topology, combinatorial enumeration, computational mathematics and advanced computational applications [1]-[19]. ,

In this paper, vectors, matrices and various measurements are described relevant to wider applications on variant construction [35]-[42] following each module as a complementary supporting documentation for the first paper “A Visual Framework of Meta Genomic Analysis on Variations of Whole SARS-CoV-2 Sequences” in this special issue.

From a genomic sequence, there are two types of outputs for three groups: 1) the A & B groups related to discrete measurements represented in either vectors or matrices; 2) the C groups related real measurements linked to entropy indices restricted in a limited line segment or a certain square region as a hierarchical organization. The main transformations and projections are described. Sample distributions and genomic index maps are illustrated.

Materials and Methods

Input on Four Meta Symbols

Since the MAS is focused on genomes, each element of input sequences is composed of four meta symbols: $\{A,C,G,T\}$ mapping to the four numbers $\{0,1,2,3\}$ in 1-1 correspondences.

The First and Second Orders of Combinations

From a combinatorial viewpoint, the first-order combination from the four symbols is composed of sixteen states as a lattice of hierarchy, as shown in Fig. 1(a).

The sixteen states $\{\emptyset, A, C, G, T, \dots, ACGT\}$ can be mapped into the sixteen numbers $\{0, 1, 2, \dots, 15\}$ to represent a 1D linear structure with 16 distinct positions.

The second-order combinations are composed of 2D 16×16 pairs of states or a 2D square with 256 positions.

Fig. 1 Four symbols in hierarchies; (a) sixteen combinations of four meta symbols in the first order on a lattice; (b) conditional probability structure

Multiple Probability Measures

When a genome contains m elements, the numbers of four meta symbols can be counted. Let $m_s, s \in \{A, C, G, T\}$ be a number of symbols s and p_s be a probability measure. We have the following equations for the multiple probability measures.

$$\begin{aligned}
 m &= m_A + m_C + m_G + m_T \\
 p_s &= \frac{m_s}{m}, s \in \{A, C, G, T\} \\
 1 &= p_A + p_C + p_G + p_T
 \end{aligned}$$

Conditional Probability Measures

For the conditional probability measures shown in Fig. 1(b), let $\tilde{p}_s$ be a conditional probability measure on s , and the following equations are satisfied.

$$\begin{aligned}
 p_0 &= \frac{m_C + m_T}{m} = p_C + p_T \\
 p_1 &= \frac{m_A + m_G}{m} = p_A + p_G \\
 1 &= p_0 + p_1 \\
 \tilde{p}_C &= \frac{m_C}{m_C + m_T} \\
 \tilde{p}_G &= \frac{m_G}{m_A + m_G}
 \end{aligned}$$

$$\begin{aligned}\tilde{p}_A &= \frac{m_A}{m_A + m_T} \\ \tilde{p}_T &= \frac{m_T}{m_A + m_T} \\ 2 &= \tilde{p}_A + \tilde{p}_C + \tilde{p}_G + \tilde{p}_T\end{aligned}$$

Under multiple and conditional probability conditions, there are ten distinct probability measures $\{p_0, p_1, p_A, p_C, p_G, p_T, \tilde{p}_A, \tilde{p}_C, \tilde{p}_G, \tilde{p}_T\}$.

Three Workflows from Input to Output

Three workflows can be identified by the type of output.

- (1) Input: Vector $\rightarrow$ Processes in Group A $\rightarrow$ Output: Vector
- (2) Input: Vector $\rightarrow$ Processes in Groups A/B $\rightarrow$ Output: Matrix
- (3) Input: Vector $\rightarrow$ Processes in Group C $\rightarrow$ Output: Real Numbers

Three Output Types

There are three distinct types of output for fifteen modules that correspond to vector, matrix and real number.

In the most three projections of the $\{A, B, C\}$ groups, both vector and matrix can be distinguished from their basic elements into three categories: genomic (four values), integer (multiple values), probability (rational values) and similarity measures (real values) to represent genomic sequences, counting, probability and similarity measures, respectively.

In the last workflows of the C groups, it uses real numbers to represent various entropy parameters as the output values.

Nine Modules in Global Projection

Nine modules $\{1DS, 2DS, RDG, PC, 1DL, MD, WP, KP, 3DV\}$ are described in three parts as input, output and process.

A_1 : 1DS similarity comparison between two sequences in m-mers.

$$\begin{aligned}\text{Input} &: N \text{ (a genome vector with } N \text{ elements)} \\ \text{Output} &: M \text{ (a similarity vector with } M \text{ elements)} \\ \text{Process} &: N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} M \xrightarrow{\text{Vector}} M \\ CF &: 16(\text{total number of selections})\end{aligned}$$

A₂: 2DS similarity comparison on one sequence in M segments.

Input : N
Output : $M \times M$ (a similarity matrix with $M \times M$ elements)
Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} M \xrightarrow{\text{Matrix}} M \times M$
CF : 32 (total number of selections)

A₃: MCP multiple and conditional probability measures of m elements in a segment.

Input : N
Output : $m + 1$ (a probability vector with $m + 1$ elements)
Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m + 1 \xrightarrow{\text{Vector}} m + 1$
CF : 10 (total number of selections)

A₄: PC protein coding of m elements in a segment.

Input : N
Output : M (a protein coding vector with M elements)
Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} M \xrightarrow{\text{Vector}} M$
CF : 16 (total number of selections)

A₅: 1DL 2D to 1D linearized transformation for m elements in a segment.

Input : N
Output : $(m + 1)^2$ (a linearized probability vector with $(m + 1)^2$ elements)
Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m + 1 \xrightarrow{\text{Matrix}} (m + 1) \times (m + 1) \xrightarrow{\text{Vector}} (m + 1)^2$
CF : 256 (total number of selections)

A₆: MD momentum distributions of probability measurements of m elements in a segment.

Input : N
Output : M (a momentum probability vector with M elements)
Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} M \xrightarrow{\text{Vector}} M$
CF : 16 (total number of selections)

A₇: WP whole/parts in n partitions.

Input : N, n

Output : $(m+1) \times (m+1)$ (a projective matrix with $(m+1) \times (m+1)$ elements)

Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m+1 \xrightarrow{\text{Matrix}} (m+1) \times (m+1)$
 $\xrightarrow{\text{Combinatorial Matrix}} (m+1) \times (m+1)$

CF : 256×2^n (total number of selections and combinations)

A₈: KP K-mer process of m elements in a segment.

Input : $N, 0 < m \leq T$

Output : $(m+1) \times (m+1)$

(a series of projective matrices with $(m+1) \times (m+1)$ elements)

Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m+1 \xrightarrow{\text{Matrix}} (m+1) \times (m+1)$

CF : $256 \times T$ (total number of selections and combinations)

A₉: 3DV transforming 2D map to 3D visualization of m elements in a segment.

Input : N

Output : $f[(m+1) \times (m+1)]$ (a visual matrix with $(m+1) \times (m+1)$ elements)

Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m+1 \xrightarrow{\text{Matrix}} (m+1) \times (m+1)$
 $\xrightarrow{\text{Visual Matrix}} f[(m+1) \times (m+1)]$

CF : 256 (total number of combinations)

Two Modules in Clustering Projection

Two modules {IDM, CDM} are described in three parts as input, output and process.

B₁: IDM $256 \times [(m+1) \times (m+1)]$ density matrix and integration.

Input : N

Output : $(256+1) \times [(m+1) \times (m+1)]$

(257 matrices each with $(m+1) \times (m+1)$ elements)

Process : $N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m+1 \xrightarrow{\text{Matrix}} (m+1) \times (m+1)$

$\xrightarrow{\text{Combination}} 256 \times [(m+1) \times (m+1)]$

$\xrightarrow{\text{Integration}} [(m+1) \times (m+1)]$

CF : 1

B₂: CDM $256 \times [(m+1) \times (m+1)]$ clustering density matrix and integration.

Input : N

$$\begin{aligned}
\text{Output} : & (256 + 1) \times [(m + 1) \times (m + 1)] \\
& \text{(a selection involved 257 matrices each with } (m + 1) \times (m + 1) \text{ elements)} \\
\text{Process} : & N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m + 1 \xrightarrow{\text{Matrix}} (m + 1) \times (m + 1) \\
& \xrightarrow{\text{Selection}} 256 \times [(m + 1) \times (m + 1)] \\
& \xrightarrow{\text{Integration}} [(m + 1) \times (m + 1)] \\
\text{CF} : & 2^{(m+1)^2} \text{ (total number of selections)}
\end{aligned}$$

Four Modules in Genomic Index Projection

Four modules {CE, IE, ME, TE} are described in three parts as input, output and process.

C_1 : CE one of 16 combinatorial entropies.

$$\begin{aligned}
\text{Input} : & N \\
\text{Output} : & 1 \text{ (a real number } \in [0, \log_2(m + 1)]) \\
\text{Process} : & N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m + 1 \xrightarrow{\text{Combination}} 16 \times (m + 1) \\
& \xrightarrow{\text{Entropy}} 16 \xrightarrow{\text{Selection}} 1 \\
\text{CF} : & 16 \text{ (total number of selections)}
\end{aligned}$$

C_2 : IE an integrated entropy.

$$\begin{aligned}
\text{Input} : & N \\
\text{Output} : & 1 \text{ (a real number } \in [0, \log_2(m + 1)]) \\
\text{Process} : & N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m + 1 \xrightarrow{\text{Combination}} 16 \times (m + 1) \\
& \xrightarrow{\text{Integration}} (m + 1) \xrightarrow{\text{Entropy}} 1 \\
\text{CF} : & 1
\end{aligned}$$

C_3 : ME a mean entropy.

$$\begin{aligned}
\text{Input} : & N \\
\text{Output} : & 1 \text{ (a real number } \in [0, \log_2(m + 1)]) \\
\text{Process} : & N \xrightarrow{\text{Segment}} m \times M \xrightarrow{\text{Measure}} m + 1 \xrightarrow{\text{Combination}} 16 \times (m + 1) \\
& \xrightarrow{\text{Entropy}} 16 \xrightarrow{\text{Mean}} 1 \\
\text{CF} : & 1
\end{aligned}$$

C_4 : TE a topological entropy[32].

Input : N
Output : 1 (a real number $\in [0, 1]$)
Process : $N \xrightarrow{\text{m-mer Segment}} m \times N \xrightarrow{\text{Counting}} \leq N \xrightarrow{\text{T-Entropy}} 1$
CF : 1

Entropy Measurements

Different from various probability measures in three projections, entropy measurements based on thermodynamics, thermostatics, statistics and statistical mechanics [20]-[34] are calculated from all M segments to collect relevant information on distributions on $(m+1)$ vectors or $(m+1)^2$ matrices. The possible schemes are {A3, A5, A7, A8, A9, B1, B2} modules to evaluate their entropy properties. For the other four modules {A1, A2, A4, A6}, all outputs are linked to M segments. It is essential to apply statistical operations to arrange intermediate outputs to satisfy the format under either $m+1$ or $(m+1)^2$ conditions, and then their entropy measurements can be evaluated further.

Let a vector Z with $(m+1)$ elements, $Z = (Z_0, Z_1, \dots, Z_j, \dots, Z_m), 0 \leq Z_j \leq M$ and $M = \sum_{j=0}^m Z_j$. Under this condition, let $P_j = \frac{Z_j}{M}$ be the j -th probability measurement, and a relevant information entropy ZE can be determined and restricted in a $[0, \log_2(m+1)]$ region.

$$\begin{aligned}
 ZE &= - \sum_{j=0}^m P_j \log_2(P_j), ZE \in \mathbf{R} \\
 1 &= \sum_{j=0}^m P_j, 0 \leq j \leq m \\
 \forall ZE &\in [0, \log_2(m+1)]
 \end{aligned}$$

Let a matrix Z' with $(m+1)^2$ elements, $Z' = (Z'_0, Z'_1, \dots, Z'_j, \dots, Z'_{(m+1)^2-1}), 0 \leq Z'_j \leq M$ and $M = \sum_{j=0}^{(m+1)^2-1} Z'_j$. Under this condition, let $P'_j = \frac{Z'_j}{M}$ be the j -th probability measurement, and a relevant information entropy $Z'E$ can be determined and restricted in a $[0, 2 \times \log_2(m+1)]$ region.

$$Z'E = - \sum_{j=0}^{(m+1)^2-1} P'_j \log_2(P'_j), Z'E \in \mathbf{R}$$

$$1 = \sum_{j=0}^{(m+1)^2-1} P'_j, 0 \leq j < (m+1)^2$$

$$\forall Z' E \in [0, 2 \times \log_2(m+1)]$$

Topological Entropy

Different from the other three entropies, topological entropy [32] is based on another set of equations. Let $P_Z(m)$ be the number of different m -length subwords (overlaps allowed) that appear in Z . Topological entropy ZE satisfies the following equation.

$$ZE = \lim_{m \rightarrow \infty} \frac{\log_4 P_Z(m)}{m}, 0 \leq ZE \leq 1$$

Under this formula, a genome Z is composed of a set of subwords $P_Z(m)$ to be transformed into a topological entropy restricted in the $[0,1]$ region.

Applying multiple entropy measurements, it is feasible to extend multi-tuples of regions to be square (two values), cube (three values) and so on to distinguish refined projections into higher dimensional geometries in detail.

Table 1 Input and output types of fifteen modules in the $\{A, B, C\}$ projections

Module	Input (Vector)	Output	Type	CF	Description
A1	N	M vector	real	16	1D similarity measures
A2	N	M^2 matrix	real	32	2D similarity measures
A3	N	$(m+1)$ vector	rational	10	ten probability measures
A4	N	M vector	integer	16	protein coding measures
A5	N	$(m+1)^2$ vector	integer	256	2D to 1D linearized measures
A6	N	M vector	integer	16	1D momentum probability
A7	N,n	$(m+1)^2$ matrix	integer	256×2^n	2D whole/parts in n portions
A8	N, T	$(m+1)^2$ matrix	integer	$256 \times T$	K-mer hierarchy of measures
A9	N	$(m+1)^2$ matrix	rational	256	2D to 3D visualization
B1	N, T	$257 \times [(m+1)^2]$ matrices	rational	1	density matrices & Integration
B2	N	$257 \times [(m+1)^2]$ matrices	rational	$2^{(m+1)^2}$	clustering density mat. & integ.
C1	N	1 scale	real	1	combinatorial entropy
C2	N	1 scale	real	1	integrated entropy
C3	N	1 scale	real	1	mean entropy
C4	N	1 scale	real	1	topological entropy

Results and Discussion

Sample Results

A list of sample results on various modules {A1, A5,A9, B1, B2, C1, C2, C3, C4 } are selected for illustrations as follows.

A1: Various curves of (A+T)% measures in multiple segments

A5: 2D to 1D linearized measures in visual maps

A9: 2D visual

A9: 3D visual

A9: 3D visual projection

B1: 257 maps on integrated density matrices

B2: 257 maps on clustering density matrices

C1: 381 SARS-CoV-2 genomes on 1D & 2D combinatorial genomic index maps

(C3,C2): 32 genomes on 2D (mean, integrated) genomic index maps

(C3,C2): 62 genomes on 2D (mean, integrated) genomic index maps

(C3,C2): 128 genomes on 2D (mean, integrated) genomic index maps

Error bar graph of topological entropy of different countries

Error bar graph of topological entropy of different countries

Error bars of topological entropy in different regions of Australia

Error bars of topological entropy in different regions of Italy

Error bars of topological entropy in different regions of France

Scatter plots of topological entropy for different countries

Scatter plots of topological entropy for different regions of the United States

Scatter plots of topological entropy in different regions of China

Scatter plots of topological entropy for different regions of Australia

Scatter plots of topological entropy for different regions of Italy

Scatter plots of topological entropy for different regions of France

Discussion

Visual distributions of the A1 module represent multiple genomes under multiple segments to show $AT - GC$ and $AT\%$ measuring curves. Checking either USA or China genomic sequences, various similarities or different parts can be identified.

Visual distributions of the A5 module show SARS-CoV-2 genomes collected from G20 regions and their refined 2D to 1D linearized measures. Checking each distribution, similarities and differences of various groups can be recognized.

In the A9 module, three sets of samples are selected from 2D maps, and 3D visual and projections are illustrated. Different coronaviruses can be distinguished by their 3D distributions.

A total of 256 maps and one integrated map of the B1 module are shown. Two sets of 256 clustering maps and integrated maps of the B2 module are selected. They are significantly different patterns with reflection and 90 degree rotational symmetries on two integrated maps.

In addition to sample results in the $\{A,B\}$ groups, further samples were selected from the C groups. Under the C1 module, 281 genomes were selected to be transformed by 2D combinatorial genomic indices. A 2D genomic index map is created. Genomes were collected from various regions in four countries: the USA, China, Australia, and Belgium. The first two rows are shown in two X-Y projections for the four countries for both horizontal and vertical projections in eight 1D histograms. The following three rows were contained in various regions that could be further separated in the relevant 1D histogram. Very rich distributions and projections could be observed.

Three sets of sample results were selected for the $\{C2, C3\}$ modules composed of horizontal axis in the C2 mean genomic index and vertical axis in the integrated genomic index. Left maps with phase-shifting effects and right maps were after further mean operations were performed. Three sets of results were contained in the $\{32, 62, 128\}$ genomes. Significant visual differences could be observed.

A total of 1337 genomes were collected and illustrated for the C4 modules. The first two rows were 1D topological genomic index maps for eight countries. The following three rows were 2D topological genomic index maps and separated results for five countries. It is interesting to see significant differences among various countries from genomic index maps with extremely different distributions from distinguished regional maps.

Conclusion

As a complementary document, this paper provides additional information to support the first paper “A Visual Framework of Meta Genomic Analysis on Variations of Whole SARS-CoV-2 Sequences” for any person who would like to understand functions of the MAS in general. From workflows of fifteen modules and their input-output data types, three types of data forms, vector, matrix and real number, are

identified for the MAS. The superpowerful capacities were built from a set of entropy parameters under thermodynamics, thermostatics and statistical mechanics. Multiple genomic index maps were shown as examples to illustrate this invariant quantitative measurement with supersymmetric properties in wider projections to provide global identifications on complicated genomes on a restricted region with infinite small scaling capacities. Further theoretical explorations and practical applications are required to be deeply explored in future activities.

Conflict Interest

No conflict of interest has been claimed.

Acknowledgements The authors would like to thank NCBI, GISAID, CNGBdb, Nextstrain and Dr. Zhigang Zhang for providing invaluable information on the newest dataset collections of SARS-CoV-2 and other coronavirus genomes to support this project working smoothly.

References

1. P. J. Cameron, *Combinatorics: Topics, Techniques, Algorithms*, Cambridge University Press, 1994.
2. J. R. Chen, *Combinatorial Mathematics*, Harbin Institute of Technology Press, 2012 (in Chinese).
3. H. W. Gould, Some Generalizations of Vandermonde's Convolution, *The American Mathematical Monthly*, Vol. 63, No.2 84-91, 1956.
4. H. W. Gould, *Combinatorial identities*, Morganton, 1972.
5. M. Hall, *Combinatorial Theory*, 2nd edition, Blaisdell, 1986.
6. L. K. Hua, *Loo-Keng Hua Selected Papers*, Springer, 1982.
7. L. K. Hua, *Selected Work of Hua Loo-Keng on Popular Sciences*, Shanghai Education Press, 1984 (in Chinese).
8. D. E. Knuth, *The Art of Computer Programming*, Vol. 1, 3rd edition, Addison-Wesley, 1998.
9. D. E. Knuth, *The Art of Computer Programming*, Vol. 4A: Combinatorial Algorithms, Part 1, Addison-Wesley, 2011.
10. F. Morgan, *Geometric Measure Theory*, 4th edition, Elsevier 2009.
11. G. Polya, R. Tarjan and D. Woods, *Notes on Introductory Combinatorics*, Birkhauser, 1983.
12. R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, 2nd edition, Cambridge University Press, 1997.
13. D. Stanton, R. Stanton and D. White, *Constructive Combinatorics*, Springer-Verlag, 1986.
14. G.Z. Tu, *Combinatorial Enumeration Methods & Applications*, Science Press, 1981 (in Chinese).
15. A. Tucker, *Applied Combinatorics*, John Wiley & Sons, 2007.
16. J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, 2nd edition, Cambridge University Press, 2001.
17. Vandermonde's identity https://en.wikipedia.org/wiki/Vandermonde's_identity
18. L. X. Wang, *An Elementary Treatise on Combinations*, Harbin Institute of Technology Press, 2012 (in Chinese).
19. L. Z. Xu, M. S. Jiang and Z. Q. Zhu, *Combinatorial Mathematics of Computation*, Shanghai Science & Technology Press, 1983 (in Chinese).

20. L D Landau, E M Lifshitz. Statistical Physics, 3rd edition, Part 1, Pergamon Press 1986.
21. HB Callen. Thermodynamics and an Introduction to Thermostatistics, 2nd Edition. John Wiley & Sons 1985.
22. W Greiner, L Neise, H Stocker. Thermodynamics and Statistical Mechanics. Springer-Verlag 1995.
23. T L Hill, Introduction to Statistical Thermodynamics, Addison-Wesley, Reading Mass. 1960.
24. R Kubo. Thermodynamics. North-Holland Pub. Co. 1968.
25. R P Freynman, Statistical Mechanics, Benjamin Reading Mass. 1972.
26. B Widom, in Fundamental Problems in Statistical Mechanics, Vol. III edi. by Cohen, Horth-Holland, 1975 1-45.
27. M E Fisher, in Critical Phenomena, Lecturer Notes in Physics, Vol. 186, ed. Hahne, Springer-Verlag 1983
28. K G Wilson, Rev., Mod. Phys. 55, 583 1983
29. J Beatte, I Oppenheim. Thermodynamics, Elsevier Scientific 1979.
30. S K Ma, Statistical Mechanics, World Scientific 1985.
31. P W Atkins, The Second Law, Scientific American Books and W H Freeman and Co. 1984.
32. Koslicki D, Topological entropy of DNA sequences, Bioinformatics, Vol.27(8) 1061-1067, 2011.
33. D Chandler. Introduction to Modern Statistical Mechanics, 1st edition. Oxford University Press Inc. 1987.
34. K Huang. Statistical Mechanics, 2nd edition John Wiley & Sons 1987.
35. Z. J. Zheng, A. Maeder, The The conjugate classification of the kernel form of the hexagonal grid, *Modern Geometric Computing for Visualization*, Springer-Verlag, 73-89, 1992.
36. Z. J. Zheng, *Conjugate transformation of regular plan lattices for binary images*, PhD Thesis, Monash University, 1994.
37. Jeffrey Z. J. Zheng, Christian H. H. Zheng, A framework to express variant and invariant functional spaces for binary logic, *Frontiers of Electrical and Electronic Engineering in China*, 5(2):163-172, Higher Educational Press and Springer-Verlag, 2010.
38. Jeffrey Z.J. Zheng, Christian H.H. Zheng and Tosiya L. Kunii. A Framework of Variant Logic Construction for Cellular Automata, *Cellular Automata - Innovative Modeling for Science and Engineering*, Dr. Alejandro Salcido (Ed.), InTech Press, 2011.
39. Jeffrey Zheng, Variant Construction from Theoretical Foundation to Applications, Springer Nature 2019 <https://www.springer.com/in/book/9789811322815>
40. Jeffrey Zheng, Variant Construction Theory and Applications, Vol. 1: Theoretical Foundation and Applications, Science Press 2020 (Chinese, Formal Publishing Soon). 郑智捷, 变值体系理论及其应用 第1册: 理论基础及其应用, 科学出版社 2020 (即将正式发行)
41. Jeffrey Zheng, ResearchGate: http://researchgate.net/pprofile/Jeffrey_Zheng
42. Jeffrey Zheng, Chris Zheng, Biometrics and Knowledge Management Information Systems, Chapter 11: Variant Construction from Theoretical Foundation to Applications, Springer Nature 2019, 193-202 https://link.springer.com/chapter/10.1007/978-981-13-2282-2_11 被斯普林格-自然杂志出版社, 选入抗击新型冠状病毒肺炎研究(Research of COVID-19)资料。推荐给 PMC 和 WHO (PubMed Central PMC and the World Health Organization WHO)。

Figures

Figure 1

Four symbols in hierarchies; (a) sixteen combinations of four meta symbols in the first order on a lattice; (b) conditional probability structure