ISBN 978-952-5726-00-8 (Print), 978-952-5726-01-5 (CD-ROM)
Proceedings of the 2009 International Symposium on Web Information Systems and Applications (WISA'09)
Nanchang, P. R. China, May 22-24, 2009, pp. 496-499

Intelligent Algorithms Analysis and Judgments Based on Thermodynamics

Ben-jun Guo, Dong-dong Chen, Jian Huang, Peng Wang
Software Engineering College,Chengdu University of Information Technology (CUIT)
Chengdu, 610225, China
guobenjun@cuit.edu.cn

***Abstract*—The article analyzes and researches thermodynamics entropy and information entropy to point out the internal relation between information and thermodynamics; it also utilizes dissipative structure theory and implicit parallelism to analyze thermodynamic essence of intelligent algorithms, then get the internal relation between thermodynamics and energy consumption of computer algorithms, implicit parallelism, intelligent algorithms, and get the conclusion that energy consumption of intelligent algorithms is lost energy. Finally, thermodynamic characteristic test of same scale algorithms and different scale algorithms is utilized to analyze and verify performance of intelligent algorithms and judge conclusion, which offers a new inspiration to research of intelligent algorithms, and offers a new train of thought to judgment on intelligent algorithm..**

***Index Terms*—thermodynamic entropy, information entropy, algorithm implicit parallelism, intelligent algorithm, energy consumption**

I. INTRODUCTION

Artificial intelligence system is essentially an information system[1,2], and the degree of intelligence is critically based on the quality of intelligent algorithms. Whoever, past researches generally focuses on algorithms structure of intelligent algorithms, etc., and biases towards accomplishment of function, so the researches on energy consumption characteristics of algorithms is relatively less.

Based on thermodynamic macro-interpretation proposed by Clausius, thermodynamic micro-interpretation proposed by Boltzmann and theory of information entropy metrics raised by Shannon, the article utilizes algorithms implicit parallelism and dissipative structure theory to analyze and judge performance of intelligent algorithms on "entropy", "energy" in thermodynamics and "information entropy" in information theory.

II. ENTROPY THEORY

Thermodynamics discusses and researches macroscopic theory of matter thermal motion, which is made up of four fundamental laws based on a great deal of experimental facts and corresponding fundamental state function, and information entropy is closely related to thermodynamic entropy.

A. Thermodynamic Entropy

Clausius proposed the concept of thermodynamic entropy on the basis of Second Law of Thermodynamics [3].Increasing or decreasing of entropy in thermodynamic system is measured by integral of temperature ratio heat in any reversible process, which is called as S, then:

$$dS = \frac{dQ}{T} \qquad (1)$$

Q presents absorbed heat, T presents temperature, and unit of entropy is $J \bullet k^{-1}$ .When system is during insulated process or is isolated, $dS \geqslant 0$, then it is principle of entropy increase: In isolated system or insulated process, entropy always increases, then the entropy defined wherein is Clausius entropy[4], or thermodynamic entropy, which is the formulation to macroscopic thermodynamic characteristic (entropy) of matter.

B. Boltzmann Entropy

In 1877, Boltzmann proposed statistical interpretation of entropy, which presents nature of entropy from the microscopic angle. Boltzmann proposed that probability of all the microscopic states in the system is equal when the whole system is in statistic equilibrium. Entropy of system is related to number of microscopic states, and Boltzmann offered a microscopic definition of entropy wherein, namely,

$$S = k \ln W_N \qquad (2)$$

In the formula, $k = 1.38 \times 10^{-23} J \bullet k^{-1}$,W_N is the number of microscopic state, and also is the measure of disorder of a system.

Boltzmann entropy exposed the relation between macroscopic state and microscopic sate, and pointed out that entropy is measure of disorder of system, or measure of uncertainty of system; Variation of entropy shows the direction of thermodynamics, and size of entropy reflects stability of state of system.

For any known isolated system, entropy only can increase rather than decrease. If entropy of system is decreased, energy must be provided from outside even if system is highly ordered and certain.

C. Information Entropy

The fundamental function of information is to eliminate uncertainty of things, and Shannon described

Ben-jun Guo,born in 1975.he is a Lecturer of Chengdu University of Information Technology. his research interests include computer software and theory.

AP-PROC-CS-09CN001

and defined information starting from the probability angle, which can be generally understood as: the acquisition of information means gathering of probability distribution in all the possibilities, and is also the probability of specific information. As the most important hardware in information society, computer generally adopts binary system, namely, each bit of data only can be 0 or 1, so probability of each is 1/2, so the necessary information amount is 1 bit to make complete judgment, or the unit of information amount. Generally speaking, in order to make complete judgment from N probabilities, the necessary information amount is $n = \log_2 N = K \ln N \ (bit)$,and probability of each possibility is $W = \frac{1}{N}$, because $\ln N = -\ln W$, the necessary information amount is $-K \ln W$ to make complete judgment, and Shannon defined it that information entropy is S, namely, $S = -K \ln W$, information entropy is lost information amount to make complete judgment, or loss of information. Information entropy H(X) of discontinuous variable is defined as

$$H(X) = -\sum p(x) \log p(x) \qquad (3)$$

The base in the above formula is 2. According to definition, information entropy is only depending on its probability distribution rather than the actual value of X.

Differential entropy h(x) of continuous variable x which takes f(x) as density function is defined as

$$h(X) = -\int_s f(x) \log f(x) dx \qquad (4)$$

In the above formula, S is the support set of the variable, and the differential entropy is only depending on probability density of variable.

If the decrease of information entropy S is equal to the increase of information amount I, then information amount is equal to negative entropy, and information entropy is equal to information amount lacked in making complete judgment.

Information entropy is measure of ordering of system. The higher the information entropy of system is, the more information amount is needed.

Ⅲ. INFORMATION THERMODYNAMICS

If the two ks both in information entropy $S = -K \ln W$ and in Boltzmann entropy $S = k \ln W_N$ are equal, then information amount can be measured by entropy. If possibility of initial state of system is P0, and possibility of final state is P1, then information I is as follows,

$$I = K \ln \frac{p_0}{p_1} = K \ln p_0 - K \ln p_1 \qquad (5)$$

Brillouin pointed out that different possibility between information can be related to number of complexion of state, thus obtaining relation between information and entropy[5]. For a system, when it is in initial state, information I_0 is 0, number of complexion is P_0, and then entropy is as follows,

$$S_0 = k \ln p_0 \qquad (6)$$

While it is in final state, information I is not equal to 0, number of complexion $P_1 < P_0$, and then entropy is as follows,

$$S_1 = k \ln p_1 \qquad (7)$$

After information is acquired in non-isolated system, decrease of number of complexion causes decrease of entropy, and the information must be provided from outside, if the entropy increases, then

$$I_1 = K(\ln p_0 - \ln p_1) = S_0 - S_1 \qquad (8)$$

$$\Delta I = -\Delta S \qquad (9)$$

Namely, information is negative entropy of system. Information entropy is equal to information amount lacked in making complete judgment, or information equal to decrease of entropy S equal to increase of negative entropy N. The following conclusion can be arrived at through comparing the formula of information entropy with Boltzmann entropy:

$$1\text{bit} = k \ln 2 \ J \bullet k^{-1} = 0.957 \times 10^{-23} J \bullet k^{-1} \qquad (10)$$

Therefore, if information stored in computer is increased by 1 bit, then its entropy must be decreased by $k \ln^2$,and entropy in environment must at least be increased by $k \ln^2$, or, at temperature of T, computer at least consumes $kT \ln^2$ to process 1 bit of information, which is the lower limit in theory of energy consumption when computer is processing.

In order to promote entropy of computer, namely, disorder, uncertainty, and mis-sequence are high, system needs more information amount and need to consume more energy.

Ⅳ .PERFORMANCE ANALYSIS OF INTELLIGENT ALGORITHMS

The accomplishment of intelligent algorithms is one of conduct of research for human, while development of artificial intelligence rests on its intelligent calculation part, namely, intelligent algorithms.

A. Dissipative structure theory

Dissipative structure theory shows that dissipative system can be converted into an ordering state of time, space or function from unordered state through continually exchanging matter and energy with the outside, which means the decrease of entropy; order and disorder always appear at the same time.

Systematic entropy variation dS is made up of entropy flow deS between system and the outside and entropy production diS in system, namely, dS=deS+diS. In the open system, dS can be plus, 0 or negative, namely, order and disorder of system are mutually transformed, namely, mutual transformation between certainty and uncertainty.

If entropy flow is less than entropy production, then dS decrease. According to statistical meaning of entropy, the part with more order, stability and certainty of the system will increase.

Dissipative structure theory is the new interpretation of entropy, full of statistical and evolutional views, and inoculates physical science, systematic science and artificial intelligence.

B. *Implicit Parallelism*

Nature of intelligent algorithms is to construct uncertain system similar to human reasoning. The system provides information, and consumes energy at the same time. The information provided realizes parallel search of algorithms on solution space, and uncertainty of algorithms causes implicit parallelism, and they are proportionate to each other[6].

Energy consumption of intelligent algorithms stems from uncertain system of intelligence, and information amount the system needed is proportionate to the uncertainty. Implicit parallelism of algorithms offers highly efficient paralleling method to make parallel search on solution space, but increases uncertainty of system, therefore reducing exactness of system algorithms.

According to relation between information entropy and thermodynamics and implicit parallelism theory[7-9], it can be concluded:

1)The higher the system entropy is, the more unordered the system is, namely, the higher the uncertainty of system is; on the contrary, the lower the system entropy is, the more ordered the system is, namely, and the lower the uncertainty of system is. It is inferred that entropy is proportionate to uncertainty.

2)The higher the entropy of computer system is, the more information is desired to provide; and vise verse. It is inferred that entropy of computer system is proportionate to desired information amount and energy consumption.

3)The higher the implicit parallelism of system is, the higher the uncertainty is, and the lower the exactness of algorithms is; the lower the implicit parallelism of system is, the higher the certainty of algorithms is, and the higher the exactness the algorithms is. It is inferred that implicit parallelism is proportionate to uncertainty, and is inversely proportional to exactness.

4)The more the information amount of system is, the lower the uncertainty is, the higher the exactness of algorithms, and the higher the intelligent of algorithms is.

Intelligent system is essentially an information system. Exactness of algorithms, consumption, and level of intelligence are all judged according to information amount. Human brain is intelligent system, while information entropy is measure of information amount, and information entropy, energy consumption are both equal to lost information amount, therefore, it can be theoretically inferred that energy consumption of intelligent algorithms can be judged by the lost energy of system.

In order to further explain corresponding relation between implicit parallelisms, entropy, desired information amount, and energy consumption, the linked structural diagram according to Figure 1 is constructed to interpret.

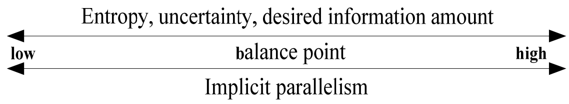

Figure 1 Linked Structural Diagram

In the diagram, the relation between intelligence and other volumes cannot be simply presented. In intelligent system, many aspects of elements should be involved to judge intelligence, such as, information amount, implicit parallelism, exactness, and energy consumption cannot be determined in linear single model, and all of these are also the content which should be pay attention to in next stage of thermodynamics, information theory and intelligent algorithms.

Ⅴ.PERFORMANCE JUDGMENT OF INTELLIGNET ALGORITHMS

The theoretical relation between thermodynamic entropy, information amount, energy consumption and intelligence are analyzed to judge energy consumption of intelligent algorithms whereon.

A. *Performance Judgment*

In the process of research, “Es (energy source)” are defined to present energy used in algorithms.

$$Es = \int_0^t [S(t) \cdot kT \ln 2 + U_0(t)] dt \qquad (11)$$

$U_0(t)$ presents internal energy dissipation caused by notoperation of computer, and it can be neglected in computer system. $S(t)$ is problem scale function, t presents time. Unit of Es is joule (J).

According to Second Law of Thermodynamics and relation between above information and thermodynamics, it can be inferred that:

$$\int_0^t S(t) dt = c \cdot H(t) \qquad (12)$$

C is proportionality constant, while Es can be simplified to be

$$Es = c \cdot H(t) \cdot kT \ln 2 \qquad (13)$$

Es presents the energy depending on loss of intelligent algorithms in desired energy. And it can be further inferred that variation of information entropy is equal to energy consumed by intelligent algorithms.

$$\Delta Es = c \cdot kT \ln 2 \ [H(t_2) - H(t_1)] \qquad (14)$$

ΔEs is equal to energy consumption of intelligent algorithms, and the conclusion is that energy source Es can be used to judge energy characteristic of intelligent algorithms

B. *Experimental Verification*

The article adopts comparing verification of same scale method and different scale method to verify energy consumption characteristic judgment of intelligent algorithms.

Genetic algorithm[10] is utilized to solve optimized search algorithms. The algorithms have implicit parallelism, and acquire exacter solution through promoting fitness and setting bigger iterative number. In each iterating process, the uncertain part is introduced by uncertainty in crossover operator (crossover probability) and uncertainty in mutation operator (mutation probability), and the involved t generation will be incorporated into the certain part of system, namely, the η(t+1) generation which will make choice, crossover, mutation is the uncertain part of system.

Table 1 shows experimental conclusion on same scale (genetic algorithms) performance judgment.

TABLE 1 Thermodynamic characteristic of genetic algorithms

Evolution time	Uncertainty	fitness	Calculating time	Implicit parallelism	Information amount
0	max	low	Short	max	few
5	bigger	lower	shorter	bigger	fewer
10	less	higher	longer	less	more
100	little	high	longer	little	many

Other intelligent algorithms are also uncertain algorithms based on probability, and expressions of information entropy of other intelligent algorithms can be inferred in the same method, so it is needless to describe hereon.

In order to further verify the energy consumption characteristics of intelligent algorithms, table 2 shows experimental conclusion of thermodynamic performance difference between genetic algorithms and other algorithms (take Simulated Annealing[11]for example) of different scale algorithms.

TABLE 2 Thermodynamic characteristic of different algorithms

algorithms	Uncertainty	Calculating time	Information amount	Information Enthalpy
Genetic algorithms	high	Same	more	Lower
Simulated Annealing	High	Same	many	Little

VI. CONCLUSION

The article arrives at the relation between information and energy thorough analyzing and discussing thermodynamic entropy and information entropy, and utilizes dissipative theory in thermodynamics and implicit parallelism in algorithms to analyze relation between “entropy”, “information entropy”, “information amount” and “energy consumption”. Besides judging intelligence of intelligent algorithms, the article expands research on energy consumption characteristics of algorithms, promotes new concern on developing direction of thermodynamics judging intelligence, and also raises new target to accomplishment of intelligent optimized algorithms.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (60702075); Sichuan Education Office Foundation(07ZA014);Development Foundation of CUIT(KYTZ200819) ; Sichuan Youth Science & Technology Foundation ; Foundation of CUIT.

REFERENCES

[1] C.H.Bennett, “Logical Reversibility of Computation”, IBM Journal of Research and Development, 1973,vol. 17, no. 6, pp. 525-532

[2] R. Landauer, “Irreversibility and Hear Generation in the Computing Process”, IBM Journal of Research and Development, 1961,Vol. 5, no.3,pp.183-191.

[3] Ling Ruiliang,Feng Jinfu. “Entropy, Quantum and Mesoscopic Quantum Phenomena”, China Science Press, 2008.1, 2~ 12.

[4] Xue Zengquan. Thermodynamics and Statistical Physics, Beijing University Press, 2004.8

[5] Feng Duan, Feng Shaotong. Entropy of the World, China Science Press, 2005.7, 232 ~ 256.

[6] Fu Chao-jiang, Zhang Wu,” An implicit parallel algorithm for nonlinear dynamic finite element analysis with overlapped domain decomposition”,Chinese Journal of Computational Mecha nics,,2006,06.pp.783-788

[7] Wang Peng. “Relation between Parallelism and Quantum uncertainty principle”, China 2008 Annual Meeting of High-Performance Center,2008.11 .

[8] Li DY. Uncertainty in Knowledge Representation, Engineering Science, 2000,2(10):73~79

[9] Li Deyi, Liu Changyu, Du Yi,and Han Xu, “Artificial Intelligence with Uncertainty”, Journal of Software,2004, vol. 15, no. 11, pp. 1583～1594

[10] Holland J H, Genetic algorithm and the optimal allocations of trials[J]. SIAM Journal of Computing. 1973,2:88-105

[11] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by Simulated Annealing, Science, vol. 220, no. 4598：671-680, May 1983