2009 International Forum on Information Technology and Applications

Decide by Information Enthalpy Based on Intelligent Algorithm

Guo Benjun, Wang Peng, Chen Dongdong, Chen Gaoyun
Software Engineering College,Chengdu University of Information Technology (CUIT), Chengdu 610225,China
guobenjun@cuit.edu.cn

ABSTRACT: In this paper we make the analysis and research about information theory and thermodynamics which includes the second law of thermodynamics, information theory, and information entropy, while pointing out the interrelation between them and coming with a new physical quantity, the information enthalpy, to communicate thermostatic and information. Finally, we get the conclusion that the change of the information enthalpy equals to the consumed energy of intelligent algorithm. We use the physical quantity to judge the performance of the intelligent algorithm and draw the relevant conclusions of the assessment experiments, which expand the intelligent algorithms based on the researches of the energy characteristics on the algorithm. In addition, the essay achieves new targets and brand-new ways of the intelligent algorithm realization and judgment.

KEYWORDS: the second law of thermodynamics; Boltzmann entropy; information theory; information entropy; intelligent algorithm ; information enthalpy

I. INTRODUCTION

The second law of thermodynamics statement, including the concept and theory of "Clausius entropy", which was proposed by Clausius in 1850, raised the discussions and researches of thermodynamics into new levels. After that Boltzmann also put forward the statistical interpretation of entropy: Boltzmann entropy, which was the micro-essence-explanation of the entropy. In 1940s, Shannon put forward the information theory and quantized it with information theory from the statistical thinking of Boltzmann. The artificial intelligence system is an information system[1] in essence, and the informational amount of which can also be measured by the information entropy.

The essay defines a new physical quantity that is information entropy through the discussion and research of the thermostatic, and uses them to connect thermostatic with information, which provides a more comprehensive judgment in the application-based realm of entropy.

II. THERMODYNAMIC THEORY

Thermodynamics is concerned with the macro theory of substance thermal movements. This article focuses on the second law of thermodynamics, which is Constant Entropy Increase Law. There are two expressions[2] of definitions on the second law of thermodynamics.

(1) Kelvin statement: it’s impossible to get heat from a single heat source to make it fully used into power without any other influence.

(2) Clausius statement: the heat can not be spread from objects in low-temperature to others in high-temperature without any other influence.

The essence of the two statements is the same: they both reveal the asymmetry of work and thermal, that is the direction of energy flowing.

Entropies are the most important physical quantities in the second law of thermodynamics and even the whole thermodynamics.

A. Clausius Entropy[3]

Entropy is a state function about the existence of thermodynamics systematic equilibrium, using the integral of heat divide with temperature of any one irreversible process from equilibrium A to equilibrium B to measure, marked as S, and then there is:

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

Q represents the heat, T represents the temperature, and the unit of entropy is $J \bullet k^{-1}$, thermodynamics entropy is also known as Clausius Entropy.

B. Boltzmann Entropy

In statistical physics, Boltzmann puts forward the statistic interpretation of the entropy. Clausius entropy is a statement of the macro-thermal characteristics of the material, which does not correspond to their accordingly micro-nature. Boltzmann submits that when the entire system exists in a statistical balance, all micro-states of the system equal to the probability of emergence, that is, the principle of equal probabilities probability. We name the macro-state number which correspond to the macro-state as the thermodynamic probability W_N of the macro-state. The largest macro-distribution of thermodynamics in the thermodynamics is the most probable distribution, which uses the Lagrange' s method of understanding multipliers to get the extreme conditions of its distribution（$\ln W_N$）. Among the huge number of micro-states in the macro system, the vast majority of the most probable distribution is in the relatively small distributions of an extremely narrow range in the vanity of the most improbable distributions, which show the same nature. The probabilistic law determines that the macro-state of the system is always evolved in the state with smaller thermodynamic probabilities to the larger ones, and the reversed actual processes can not happen. As to this, Boltzmann makes a micro definition of the entropy, that is

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

K represents the Boltzmann constant, $k = 1.38 \times 10^{-23} J \bullet k^{-1}$.

978-0-7695-3600-2/09 $25.00 © 2009 IEEE
DOI 10.1109/IFITA.2009.112

C. The Principle of Constant Entropy Increase

The principle of constant entropy increase means that in the irreversible process, the entropy of system is always larger and larger, which is a rule cannot be changed:for the isolation system without any external intervention, the probabilistic law decides that the entropy can only spontaneously increase but decrease, and the entropy gets the maximum when it reaches a balance.

D. Entropy Calculation

The process of entropy calculation is also related to some characteristics entropy:

(1) Entropy is an extensive quantity, the total entropy in the system equals to the total amount of various parts.

(2) The change of system entropy can be divided into two parts in forms: one is caused by the interaction of system and the outside environment, that is, the result of flow-in and flow-out of substance and energy. This part can be called Entropy flow, marks as $d_e s$. The other part is caused by the inner irreversible process of the system. This part can be called entropy production item, marks as $d_i s$. Then there is $ds = d_e s + d_i s$. By the definition we can see, $d_e s$ does not have certain notation, but will be greater than 0 all the time, that is $d_i s \geq 0$.

Entropy is not a conserved quantity. The change of total entropy in a system can be recorded:

$$\frac{ds}{dt} = \frac{d}{dt}\int_v s dv = \int_v \frac{\partial s}{\partial t} dv = -\int_{\sum} d\sum n \bullet j_s + \int_v dV\delta \qquad (3)$$

j_s stands the entropy exchanging rate through unit area; δ stands the entropy producing rate in unit volume.

III. INFORMATION ENTROPY

The contemporary information theory mainly focuses on the research of information amount. The information research is the description and definition of information from the view of probability with the application of Shannon; it could be summarized as: the obtainment of information means the focus of probability distribution in a variety of possibility. The computer commonly uses binary data of 0 or 1, only two states; treats the information amount fully determined in the two possibilities as 1 bit, that is, the unit of information amount. In general, the information amount in a full decision in N kinds of possibilities, which can be represents as:

$$n = \log_2 N = K \ln N \ (bit), \quad K = \frac{1}{\ln 2} bit = 1.4427 bit$$

On the condition that there is no information under N kinds of possibilities, it is assumed that each of the possibility of the probable emergence is w=1/N, as $\ln N = -\ln W$, the information amount in a full decision from N kinds of possibilities is $-K \ln W$, Shannon defines it as information entropy S[4], and then $S = -K \ln W$, Information entropy equals to the lacking information amount for a full decision, while meaning the information defects. The information entropy H(X) of discrete random variable X is defined as:

$$H(X) = -\sum p(x) \log p(x) \qquad (4)$$

The information entropy is in fact a random variable X of the pan-distribution function. It does not depend on the actual value of X, but only its probable distribution.

The information entropy h(X) of continuous random variable X that relies on the density function f(x) is defined as:

$$h(X) = -\int_s f(x) \log f(x) dx \qquad (5)$$

Here S is the set to support the random variable; differential entropy only depends on the density of random variable.

The reduction of information entropy S equals to the increase of information amount. So if the information amount is negative, then the information entropy is equal to the lack of information to an absolute decision.

IV. THE RELATIONSHIP BETWEEN INFORMATION AND THERMODYNAMICS

Order the K in the Shannon information entropy formula $S = -K \ln W$ equals to Boltzmann constant K, then the information amount could be measured by information entropy. If the possibility of the initial state existence in an information system is P_0 and that of terminal state is $P1$, then the information I:

$$I = K \ln \frac{p_0}{p_1} = K \ln p_0 - K \ln p_1 \qquad (6)$$

Brillouin pointed out that the different possibilities of information could be connected with the postural number of complexion[5], in order to get the relationship between information and entropy. Supposed a system under the initial state, information I_0 =0, number of complexion is P_0, and the entropy is:

$$S_0 = k \ln p_0 \qquad (7)$$

However in the terminal state, information $I_1 \neq 0$, number of complexion $P_1 < P_0$, entropy

$$S_1 = k \ln p_1 \qquad (8)$$

Obviously, under the concerning circumstances, the system is not isolated. When the information is obtained the number of complexion decreased, which leading the reduction of entropy. However, the information must be supplied by the outside environment, then its entropy increase, it is:

$$I_1 = K(\ln p_0 - \ln p_1) = S_0 - S_1 \qquad (9)$$

As a result, the information is equal to a negative quantity of the total entropy in a physical system.

The effects of information cause the reduction of entropy, so the information taken as the negative item of system entropy.The information indicates that the reduction of entropy S and the increase of negative entropy N. Entropy is referred to a large number of particles of state, which is a macro-volume, comparing with the information process of computer that only involves multiple choices among a small number of states (0 and 1). However, they are essentially unified.

When the probability of various possibilities is not equal, the information entropy is just equal to the lacking information for a full determination. In this process, the decrease of the information entropy is equivalent to the increase of information amount $\Delta I = -\Delta S$, that is, the information equals to the negative entropy.

The information entropy formula $S = -K \ln W$ is very similar to the Boltzmann entropy formula $S = k \ln W_N$, the only difference between them is that the unit of entropy in thermodynamics is $J \bullet k^{-1}$ and $k = 1.38 \times 10^{-23} J \bullet k^{-1}$, and the unit of that in information theory is bit. From the comparison we can get the important conclusion:

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

By (10) formula and the principle of entropy increase, we can see that if the information stored in a computer storage increased 1 bit, its entropy will reduce $k \ln 2$. So the entropy in the environment should increase so much at least. That is, the required energy when the computer process I bit information under the temperature T is $k \ln 2$ at least, which should be the minimum theoretical energy expenditure for a computer information processing.

V. INTELLIGENT ALGORITHM AND THERMODYNAMICS

The realization of the artificial intelligence is essentially an information system. Priori knowledge of the intelligence algorithm [6, 7] can be expected in the right direction of the search, but uncertainty information is provided by the system to achieve the parallel searching of solution space to algorithm. The uncertainty and the implicity of the algorithm are in direct proportion with each other[8, 9]. With the usage of this characteristic we can construct the following "Parallel Judge Pistons", which can be described as:

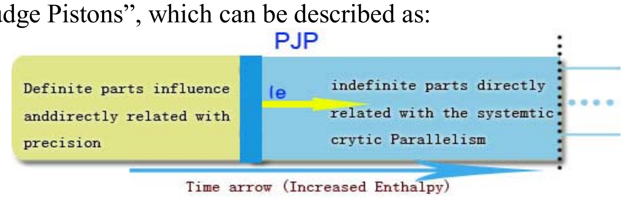

Figure 1: PJP (parallel Judge Pistons)

In order to get a better research, we define a new physical quantity"Ie (Information enthalpy)" in form:

$$Ie = \int_0^t [S(t) \cdot kT \ln 2] dt \quad (11)$$

$S(t)$ is a problem size function, the unit of Ie is J., the system piston moves to right automatically with the influence of Ie as it is shown in the figure 1.

Before the implementation of intelligent computing, the uncertainty of the system is in maximum. At this moment the information amount required is in top, so is the energy required. As the time going on, and the operation carrying out, the certain parts are in a growing number and the accuracy of computing becomes higher. Meanwhile, the result of the uncertain parts become less and less, and the implicit parallelism of the system reduces. For the convenience of research we define another physical quantity "intelligent enthalpy" H_i. It is the quotient of information system enthalpy and the time of algorithm implementation that expresses the average energy consumption of algorithm, that is:

$$H_i = \frac{Ie}{\Delta t} = \frac{\int_0^t [S(t) \cdot kT \ln 2] dt}{\Delta t} \quad (12)$$

By combining the second law of thermodynamics and the relationship between the information and thermodynamics we can get:

$$\int_0^t S(t) dt = c \cdot H(t) \quad (13)$$

The $H(t)$ here is the information entropy at t time (necessarily use time indicators), c for the constant ratio.

To combine formula (12) with (13), we can get the general form of information enthalpy:

$$Ie = c \cdot H(t) \cdot kT \ln 2 \quad (14)$$

The physical meaning of the information enthalpy can be clarified from its normal form, which is that the information enthalpy is the part of the intelligent algorithm lacking in the needed energy in a system. Furthermore, it is also concluded that the change of information enthalpy is the energy which is consumed in intelligent algorithm.

$$\Delta Ie = c \cdot kT \ln 2\ [H(t_2) - H(t_1)] \quad (15)$$

$$\frac{dIe}{dH} = c \cdot kT \ln 2 \quad (16)$$

Here the energy consumption should be ΔIe.

We can use information enthalpy to determine the energy consumption characteristics of intelligent algorithm. Then we will use the genetic algorithms as the examples to demonstrate the above conclusions.

In the genetic algorithm, algorithms are designed to achieve optimal solutions. Only by improving theirs fitness and setting up a greater number of iterations can we obtain a more accurate solution. The uncertain parts in every iterative process are caused by the crossover of uncertainty and the mutation of uncertainty. The evolving generation t will be involved in the certain parts of the system, and the generation η (t+1), which is going to be chosen, crossed and mutated, is the uncertain part of system. If the number of initial iteration of the system is M and that of the evolution is t, then the information entropy of the system in the form is:

$$H=-\sum_{x} p(x)\log_2 p(x)=-\sum_{i=1}^{M-t} p(t+i)\log_2 \frac{1}{p(t+i)} \quad (17)$$

Table 1 offers the experimental conclusion for the evaluation of genetic algorithm performance based on information enthalpy.

TABLE 1: GA thermodynamic properties

Evolution Algebra	**Information Enthalpy**	**Algorithmic Uncertainty**	**Calculate Time**	**Algorithmic Implicit parallelism**
0	High	Max	Short	Max
5	Higher	Bigger	Shorter	Bigger
10	Lower	Smaller	Longer	Smaller
100	Low	Small	Long	Small

VI. CONCLUSION

This paper is based on the discussion and research of information theory and thermodynamics to get the conclusion that the information is inseparable to energy, and put forward a new physical quantity "information enthalpy" to connect them in order to research the intelligent algorithm more conveniently. From the information enthalpy formula we can see that the energy consumption of intelligent algorithm does not decline without limitation. In this essay, we provide a new way of thinking to the research on the computer intelligence (artificial intelligence) and a new train of thought for the evaluation of intelligent algorithms.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (60702075); Sichuan Education Office Foundation(07ZA014);Development Foundation of CUIT(CSRF200701 and KYTZ200819).

REFERENCES

[1] Wang Peng, Parallel Computing:Application and Practice, China Machine Press, 2008.11.

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

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

[4] (U.S.)Thomas M. Cover & Joy A. Thomas Elements of Information Theory, China Machine Press, 2005.5,134 ~ 165.

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

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

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

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

[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.