Calculation of the Minimum Time Complexity Based on Information Entropy

Xue Wu

Department of Microelectronic and Nanoelectronic Engineering,
Room 130A, Zijing1#Building, Tsinghua University, Beijing, P.R. China
wuxiqh@yahoo.cn

Abstract. In order to find out the limiting speed of solving a specific problem using computer, this essay provides a method based on the entropy of information. The relationship between the minimum time complexity and the information entropy change is illustrated. Several examples are served as evidence of such connection. Meanwhile some notices of modeling these problems are proposed. Finally, the nature of solving problems with computer programs is disclosed to support the theory and a redefinition of the information entropy in this field is proposed. This will develop a new field of science.

Keywords: information entropy, time complexity, algorithm, entropy change.

1 Introduction

As new approaches to algorithm optimization become more and more popular in researches, it is a matter of greater urgency to determine whether an algorithm of handling a specific problem reaches the limiting speed or not instead of trying to improve an algorithm continuously. Such demand arouses a technique to discover the boundary of the velocity, the minimum time complexity of a program, which is based on two assumptions. These assumptions consequently give birth to the relevance between information entropy change and the so-called minimum time complexity.

Information is an abstract conception. There was not a widely acceptable measurement of information quantification until the father of information theory C. E. Shannon provided a novel conception of information entropy in 1948 [1], [2]. Shannon first associated probability with information redundance in mathematic language [1]. His discovery made a great contribution to the field of communication, meanwhile it also left clues to the consistency between information entropy change and the nature of figuring out an issue using computer [3], [4], [5]. This directly led to the establishment of the first assumption.

Since computers were invented, the amount of information which is generated by an operation in a computer program has remained unknown to most people. The significance of operations' productivity has even been ignored. However it is necessary to concentrate on the efficiency of an operation with the purpose of building the bridge between information entropy change and the minimum time complexity. Therefore the second assumption arises.

N. Meghanathan et al. (Eds.): CCSIT 2012, Part II, LNICST 85, pp. 376–385, 2012.

After the proposition of these two fundamental assumptions, there comes the great need of evidence to support this theory. Just then two examples appear and become the pillars of the theory. Looking for the maximum value and sorting a group of numbers are the names of these two problems. The results of the calculation based on information entropy for these two issues in some special models are the same with the time complexities of the known fastest programs. This fact motivates a demand for notices of modeling these problems. With the help of the notices, the nature of solving problems using computer programs is discovered to testify the theory and a redefinition of the information entropy is proposed. This event heralds a new field of science to be developed [5].

2 Two Basic Assumptions

During the information transmission and storage process, there is a part of code that does not express the substance of the information. The amount of such code is called information redundance. Moreover making people realize some unaquainted knowledge is able to be regarded as giving people some substantial information [6]. Therefore the procedure is exactly the change of information redundance. According to the relationship between information redundance and information entropy, the first assumption that the nature of figuring out a problem is the same as a change of information entropy is set up.

No matter how complex an operation is, it is composed of some simple ones. Operations like equation, being greater than and being smaller than are typical instances of these basic operations. In consideration of them, each elementary operation is able to figure out a question like whether a statement is true or false. The amount of information that consists in such questions is regarded as 1 bit. Therefore the fact that every basic operation generates 1 bit information is presented as the second assumption.

3 Theory Description and Modeling of Two Examples

The fundamental role of information is to eliminate people's uncertainties of matters around them. The information entropy is used to measuring the degree of these uncertainties [6]. Shannon's conception of the entropy is based on the following equation:

$$H(x) = -\sum_{i=1}^{n} P(x_i)\log_2(P(x_i)) \tag{1}$$

In the equation, H(x) is the information entropy, $P(x_i)$ is the probability of the incident that x equals to x_i [7]. In line with the two assumptions a theory that the minimum time complexity is the same with the change of information entropy is brought out. It can be expressed in the following form:

$$\text{Time complexity} = \Delta H(x) = H(x) - H_0(x) \tag{2}$$

$H(x)$ is the initial entropy. $H_0(x)$ is the final entropy [8]. With regard to a specific question, the initial entropy is fetched through modeling of this problem [1]. The modeling criterion has not been clearly clarified, however an imprecise standard that the sequence which expresses the final relationships among data or belongs to the same sort of sequence with the final result should be regarded as the state x_i is proposed. This rule will be interpreted by two examples in detail later on. When a problem is figured out, the information which people want to know is determined. In other words, the sequence is eventually decided. Therefore the probability of the situation that this sequence appears is 1 [9]. According to (1), the final entropy of an issue is obtained.

$$\begin{aligned} H_0(x) &= \log_2(1) \\ &= 0 \end{aligned} \tag{3}$$

Backed by the modeling standard and the value of the ultimate information entropy, the minimum time complexities of looking for the maximum value and sorting a group of numbers are easy to discuss.

3.1 Modeling of Seeking the Maximum Value

In order to make a model of looking for the maximum value, the form of the sequence that is able to expresses the characteristics of the final result should be decided first. We can abstract this problem into the following structure:

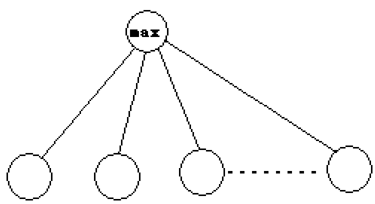

Fig. 1. Abstract structure of seeking the maximum

The top element in Fig. 1 represents the maximum value and the other elements below it are the rest of the numbers. The lines between them are called keys. They denote the size relationships between the maximum value and the others. Each size relationship has two possible conditions: “Greater” or “Not greater”. If each key’s condition is determined as “Greater”, then the top element turns out to be the maximum. This fact indicates that these keys convey the final result. Therefore the serial of keys can be regarded as the sequence that we want.

There are n-1 keys in this sequence. According to the statement that each key has two possible situations, there are 2^{n-1} different sequences in all. Every sequence has a probability of $(0.5)^{n-1}$ to be the final sequence. Use (1), we can obtain the initial information entropy.

$$\begin{aligned} H(x) &= -\sum_{i=1}^{2^{n-1}} 0.5^{n-1} \log_2(0.5^{n-1}) \\ &= n-1 \end{aligned} \tag{4}$$

As $H_0(x)$ equals to 0, with the help of (2) the minimum time complexity is got from the result of the calculation.

$$\text{Time complexity} = \Delta H(x) = n-1 \tag{5}$$

Therefore the calculated minimum time complexity is O(n). According to the survey of various algorithms, the time complexity of the fastest program which is able to solve the problem is also O(n). These two time complexities are the same in such model. This indicates that the theory of the consistency between the information entropy change and the minimum time complexity of the computer program is proved in this issue.

3.2 Modeling of Sorting a Group of Numbers

Being similar to the circumstances of seeking the maximum value, another type of sequence needs to be decided before we calculate the minimum time complexity of sorting a group of numbers. Inspired by the idea of seeking the maximum value, keys between every two elements should be established in order to express the final relationships among data. Just as the circumstances in solving the problem, looking for the maximum, each key has two possible situations. If we decide the value of every key, the correct sequence will be presented. With the thinking as before, the sequence of these keys is considered what we want. Meanwhile, the number of the sequences is considered to be $2^{0.5n(n-1)}$ with the help of the permutation and combination theory. According to this method to do so, we will get a result like this:

$$H(x) = 0.5n(n-1) \tag{6}$$

Therefore, the minimum time complexity is O(n^2). This time complexity is the same with that of Bubble sort and insertion sort [10]. These methods are fast enough, however, just as we know, the fastest algorithm is merge sort the time complexity of which is able to reach O($n\log_2(n)$) [10]. O($n\log_2(n)$) is much smaller than O(n^2), this fact indicates that O(n^2) is not the minimum time complexity and O($n\log_2(n)$) takes place of it. As a result, there may be a contradiction in this theory. Nevertheless, if we consider this issue more carefully, we will discover that we have ignored some important point, the independence of the keys. For example, if there are only three numbers, all situations which are based on the theory of permutation and combination are shown in the following figures:

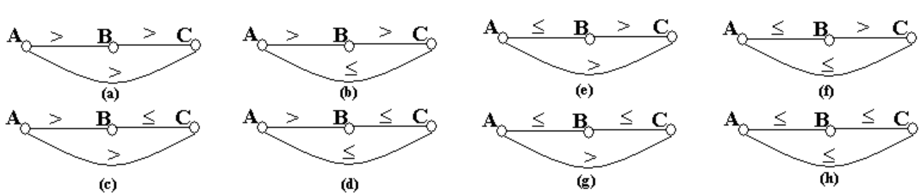

Fig. 2. Situations of three elements

In reality, not all of the eight situations are able to occur. Fig. 2(b) and Fig. 2(g) are two examples of these circumstances. Taking account of the size relationships among them, there are contradictions in the two figures cited above. Thinking about Fig. 2(b) in details, we'll get the following information: A is greater than B; B is greater than C; A is not greater than C. According to the first two points, A should be greater than C, however Fig. 2(b) provides an opposite relation. Therefore the situation in Fig. 2(b) does not exist. The same conclusion is drawn from Fig. 2(g).

As the total number of data grows, more and more situations appear and we directly regarded them as legal ones, but in fact they should not be included in the collection of the sequences. If we exclude them, the total number of the situations will correspondingly reduce to $n!$. It is easier to think in another way in order to explain why the number equals to $n!$ after reduction. When the serial of keys are decided, one sequence of the numbers is determined. Meanwhile one sequence of the numbers indicates one sequence of "Greater" or "Not greater". This is true, because the sequence that we choose is able to express the characteristic of the final result and the sequence of the numbers is exactly the result in the problem of sorting a group of numbers. Therefore, the sequence of the numbers is capable of being the state x_i. Based on the permutation and combination theory again, the number of the states is $n!$. For the reason that every state has the same probability to be the ultimate one, (1) and (2) provide us the following information:

$$\text{Time complexity} = \log_2(n!) \tag{7}$$

Time complexity is an approximation of a program's speed. Moreover, the number of the elements should be large enough in order to make the approximation meanful. In the case that n is a large number, there is a classical approximaiton.

$$\ln(n!) = n\ln(n) - n \tag{8}$$

Based on (7), we can get the following result:

$$\begin{aligned} \log_2(n!) &\approx n\log_2(n) - n\log_2(e) \\ &\approx n\log_2(n) \end{aligned} \tag{9}$$

Therefore, the minimum time complexity is O($n\log_2(n)$). The result is the same with that of the merge sort [10]. This fact confirms the correctness of the theory again.

4 Notices of Modeling Problems

Even though there is not a precise criterion of modeling problems, several notices should be clarified in order to get the correct result.

4.1 Relationship between Final Result and the Sequence

In the problem of sorting a group of numbers, we are able to choose the permutation of the numbers as the sequence because the final result is one of the permutations.

This fact proposes a new decription of the criterion that the state and the final result belong to the same type of sequence. Meanwhile the characteristic that the sequence should express the final relationships among data is eliminated. However, such modification of the rule is not correct.

Considering the problem of seeking the maximum value, if we regard the final result as the sequence, then an element is considered as the sequence and the number of the states will be n. According to (1), we are easy to find out that the minimum time complexity is $\log_2(n)$ in this model. As we know, the time complexity of solving this problem is not able to decline to this amount. Therefore the final result can not serve as the state in all problems.

However being aware of which problem's final result is capacable of being the sequence is not enough, we need to discover the reason why some of them are able to be the sequence and the others are not.

In the problems of sorting a group of numbers, the final result is a serial of numbers in order. We regard this type of sequence as the state and we are able to calculate the minimum time complexity which conforms to that of the fastest program. The reason why this consistency is established can be explained in the following way. The result of sorting a group of numbers is a serial of numbers in order. The only feature of the result is the order. Therefore the final result itself is able to express its characteristic. We call this trait information independence. When the problem is solved, the information entropy reduces to 0. This means that the sequence we choose is decided. Fortunately, the sequence we determine is the final result and because of its information independence the final result contains all the final relationships among data. Thus, in this problem we can regard the final result as the sequence.

Nevertheless, the circumstances of seeking the maximum value are different from those of sorting a group of numbers. The final result of seeking the maximum value is the maximum number. Even though the maximum number shows us its magnitude, it is not able to convey its characteristic, being the maximum. Therefore, the consequence of this problem is not information independent and this fact leads to the invalidity of regarding the final result as the sequence.

In a nutshell, the sequence which is considered to be the state should express the final relationships among data. Only if the final result is information independent, is it possible to substitude the final result for the sequence in order to simplify the state calculation.

4.2 Final Information Entropy and the Sequence

In order to determine which sequence is able to be taken as the state, the imprecise criterion is not enough, because there are various sequences which are capable of representing the final relationships among data. For instance, in the problem of looking for the maximum value, besides the sequence we choose, a sequence that contains the relationships between each two numbers accords with the criterion too. Therefore, another factor should be taken into account for the purpose of excluding the extra sequences.

Recalling the discussion in Section 3, when a problem is solved, the information entropy declines to 0. This means that the sequence is determined. However this fact will not happen, if we regard the sequence which expresses the relationships between each two numbers in the problem of seeking the maximum value. When the maximum value is located, there are only *n*-1 keys determined in the sequence and the rest of the keys remains unknown. Therefore, the entropy of information is not 0 in this case and the magnitude of the information entropy should be calculated again. This will bring in more work. Not only do we need to calculate both initial and final information entropies, we also need to determine the locations of the *n*-1 keys in the sequence.

Whereas, thing is not that easy as it looks like. The locations of the *n*-1 keys are able to provide several different states. Meanwhile, the sequence itself has several states too. It seems right to take these two types of states into account in order to find out the initial and final number of the states. But these two types of states are not totally different, there does exist an overlap between them. Therefore, if we consider in this way, modeling this problem will be more complex and a great amount of extra work should be done.

As a result, it is a good way to model problems with the sequence that makes the final information entropy decline to zero. We call this type of sequence 0-type sequence. Choosing the 0-type sequence as the state is a routine way to model problems. However in some problems, the 0-type sequence is not easy to find. This arouses another sequence to replace it. The ideas above remind us that we are able to regard some sequences that are in accord with the imprecise criterion as the state. If we are easier to handle the complex relationship between those two types of states than to establish the 0-type sequence, it will be a better way to select this sequence instead of the 0-type sequence.

4.3 Equation Simplification in Specific Circumstances

While calculating the time complexity of the two problems in Section 3, we can find out that the equation is able to be simplified. Because the possibilities of all states are the same and the final information entropy is zero, the equation can be expressed in the following form (*n* is the number of the states):

$$\text{Time complexity} = \log_2(n) \tag{10}$$

This simplified equation discloses another characteristic of the states in the models of problems. That is the equality of the states' possibilities. This feature is a factor that we need to consider when we are going to select the sequence and it is also an inevitable result because the sequence satifies the imprecise criterion. The necessity can be proved by the following evidence.

If some states' probabilities of being the final sequence are different from others, this will indicates that these states have different numbers of sub-states. Since each state is able to be the final sequence, meanwhile the sequence and the result is one to one, therefore if a state with one sub-state can be the final sequence and a state with two sub-states is also capable of being the final sequence, the two results which they provide belong to different types. This means that the format of the final result is not determined. However in a specific problem the format of the final result is settled. Thus, it is not possible to make a state with different probability.

5 Nature of Solving Issues with Computer Programs

In Section 4, the equation to calculate the time complexity is simplified. According to (10), there is a logarithmic relationship between the minimum time complexity and the number of the states.

As everyone knows, the time complexity of the fastest program for searching, bisearch also has a logarithmic relationship with the number of the elements. However this coincidence does not happen occasionally, a connection does exist between searching and solving a problem with computer program.

Think about it theoretically. In Section 3, solving problems using computer programs is interpreted as determining which state is able to be the final state. As the theory says, at the beginning of solving problems, there are a number of states and we don't know which state is capable of being the final one. These states act as a list and our target is to find out which state is in accord with the characteristics of the final result. The list is named as lookup table of problems and the final result is called searching target. Therefore, solving a problem with computer programs is transformed into a problem of searching. The nature of resolving issues turns out to be raveling out a problem of searching.

Since the minimum time complexity of searching is $\log_2(n)$ (n is the number of the elements), so the minimum time complexity of any problem is able to reach $\log_2(n)$ (n is the number of the states).

6 Redefinition of Information Entropy

In Section 5, no matter what the problem is like, it is able to be considered as a searching problem when we want to solve it by computer programs. Hence the minimum time complexity is related to the number of comparisons which are required to find the searching target. Based on the theory, if the connection between the minimum time complexity and the information entropy is valid, there will be a relationship between information entropy and the corresponding searching problem. Considering the fact that the information entropy has a similar logarithmic relationship to that of the searching problem, a redefinition of the information entropy is proposed in the field of settling issues.

In the redefinition of the information entropy, it is defined as the minimum number of comparisons, which computers need to execute in order to find the searching target in the lookup table of the problem in the worst case. This definition focuses on using computer programs to solve problems. Meanwhile it is also in accord with the original definition proposed by Shannon based on the calculation in Section 4. Therefore the new definition is a special case of Shannon's definition.

Computer is a discrete binary system. It is able to resolve problems with definite results. According to the discussion in Section 4, the states of the problem are equiprobable. Thus, the equation of the information entropy can be reduced to the simple logarithmic form while considering such problems. This is the reason why we replace the original definition by the new one in this field. Moreover, this definition helps us to think about this type of issues more directly and conveniently.

7 Conclusion

As the result of the assumption that the process of solving a problem is the same as making a change in the information entropy and the postulation that an operation in computer program is able to generate 1 bit information, the theory that the minimum time complexity of settling a specific issue with computer programs equals to the information entropy change in the model of the problem with special characteristics is established.

Seeking the maximum value and sorting a group of numbers are served as evidence to support this theory. The coincidence that these two time complexities equal to the results of the calculation backs the validity of the principle. Meanwhile, (2) is proposed to calculate the minimum time complexity.

While we are calculating the minimum time complexity based on the information entropy, a pivotal problem of how to modeling issues arises. Thus, several notices are put forward in order to make the process of modeling problems more complete.

The first notice is about the imprecise criterion. This notice remind us that not all the problems are able to regard their final results as the sequences. If and only if the final result of the problem is information independent, is it correct to choose the final result as the sequence. In other words, the only criterion of selecting the sequence is that the sequence must be able to convey all the final relationships among data and taking the final result as the sequence is only a simplification of a special case.

Secondly, while selecting the sequence, there are a lot of sequences that are in accord with the imprecise criterion. This notice acts as a guide for us to choose the most befitting sequence among them. In the notice, the sequences are separated into two types. One is the sequence with zero final information entropy, which is called 0-type sequence; the other is the sequence with non-zero final information entropy. While calculating the number of the states, the circumstances of the sequence with non-zero final information entropy are very complex, but the cases of the 0-type sequence are easily figured out. However, the 0-type sequence is not usually established easily. If we can consider the complex circumstances of the sequence with non-zero final information entropy clearly, we will be able to take such sequence instead of the 0-type sequence. Therefore, the notice indicates that we should choose the 0-type sequence to make the calculation easier in general cases and we can select the sequence with non-zero final information entropy in some special cases, if we are able to think about the circumstances clearly.

An important characteristic of the states in the models of the problems is clarified in the final notice. This characteristic is described as the equiprobability of the states. In the analysis, the characteristic is proved to be necessary, because the result of the problem has a certain format. Meanwhile the equiprobability also leads to a simplification of the equation to calculate the minimum time complexity.

After the simplification of the equation is made, the similarity between the calculated result and the minimum time complexity of the searching problem is discovered. Such similarity arouses a further consideration of the relationship between the searching problem and the nature of solving problems with computer programs. Since the minimum time complexity of searching has the same form with that of the simplified information entropy and there are lookup table and the searching target in the model of the problem, so the nature of solving problems with computer programs is regarded as resolving searching problems.

With the help of the discovery that the nature of solving issues with computer programs is no more than raveling out searching problems, the information entropy is redefined as the minimum number of comparisons, which computers need to execute in order to find the searching target in the lookup table of the problem in the worst case. This redefinition is in accord with the original definition made by Shannon.

In summary, all the conclusions above herald a new field of science to be developed.

References

1. Gang, C.: Shannon Information Model in E-commerce Information Analysis. In: International Joint Conference on Artificial Intelligence, pp. 580–583 (2009)
2. Verdú, S.: Fifty Years of Shannon Theory. IEEE Transactions on Information Theory 44(6), 2057–2078 (1998)
3. Wang, Y.: Analyses on Limitations of Information Theory. In: International Conference on Artificial Intelligence and Computational Intelligence, pp. 85–88 (2009)
4. Shannon, C.E.: A Mathematical Theory of Communication. Bell System Technical Journal 27, 379–429, 623–656 (1948)
5. Shannon, C.E.: The Bandwagon. IEEE Transactions on Information Theory 2, 3 (1956)
6. Hatfield, A.J., Hipel, K.W.: Understanding and Managing Uncertainty and Information. In: IEEE International Conference on System, Man and Cybernetics, pp. 1007–1012 (1999)
7. Zhao, M., Ye, X., Han, K., Li, Y.: Research on Digital Image Edge Detection with Local Entropy and Fuzzy Entropy Algorithms. In: IEEE International Conference on Information and Automation (ICIA), pp. 2477–2482 (2010)
8. Xie, X., Zang, X.-Y., Guan, Z.-L.: Analytic Demonstration on the Irrationality of Negative Entropy Principle. In: International Conference on Management Science and Engineering, pp. 109–115 (2008)
9. Wang, X., Ge, Y., Wang, X.: Research on Manthematical Theory of Information Acquisition. In: Proceedings of International Conference on Information Acquisition, pp. 59–70 (2004)
10. Iqbal, S.Z., Gull, H., Muzaffar, A.W.: A New Friends Sort Algorithm. In: 2nd IEEE International Conference on Computer Science and Information Technology, pp. 326–329 (2009)