Eur. Phys. J. B **48**, 575–581 (2005)
DOI: 10.1140/epjb/e2006-00010-8

THE EUROPEAN PHYSICAL JOURNAL B

Evolutionary model with intelligence and knowledge

Q.H. Pan[2,3], B.L. Yu[1,3], and M.F. He[1,3,a]

[1] Department of Applied Mathematics, Dalian University of Technology, Dalian, P.R. China
[2] Department of Applied Physics, Dalian University of Technology, Dalian, P.R. China
[3] Institute of University Students' Innovation, Dalian University of Technology, Dalian, P.R. China

Received 25 June 2005 / Received in final form 30 October 2005
Published online 19 January 2006 –

Abstract. An evolutionary model based on Bit-String with intelligence and learning is constructed. Each individual is represented by five bit strings. Four of them relate to genes and the other denotes the knowledge from learning. The four strings relative to genes will be divided into two parts — Bit-String A and Bit-String B. Bit-String A denotes the health of an individual, at the same time Bit-String B can describe intelligence. For an individual, the cross reproduction method is used in this model. After that we explain how knowledge is represented in our model. The probability of learning is affected by intelligence. In order to study how the accumulated knowledge influences the survival process by the effect of food and space restrictions, we modify the Verhulst factor. Then, we present the results of our simulations and discuss the evolution of population, intelligence and knowledge respectively. In addition, an equation to calculate the intelligence quotient is given based on intelligence and knowledge. We discuss the distribution of intelligence quotient.

PACS. 87.23.Cc Population dynamics and ecological pattern formation – 89.75.Da Systems obeying scaling laws – 05.10.Ln Monte Carlo methods

1 Introduction

Recently, many genes with different kinds of function have been found. Many of these are related to health. Enrique Gonzalez, et al. discuss the influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility [1]. And Hannes Lohi, et al. performed a similar study in canine epilepsy [2]. There are also some genes related to intelligence. Virginie S. Vervoort, et al. found a gene named AGTR2 [3]. Their findings indicate a role for AGTR2 in brain development and cognitive function.

Penna presents a simple model for biological aging based on Bit-Strings [4]. The model works under the effect of the Verhulst factor, mutations, death by genetic diseases or age and a minimum reproduction age. The sexual Penna model introduced by Stauffer et al. [5,6] corresponds to a reproductive regime of diploid organisms, the population being divided into males and females. Bustillos AT and de Oliveira PM represent a process of learning using Bit-Strings based on the Penna model [7]. They study how knowledge is accumulated during life and its influence over the genetic pool of a population after many generations. But the knowledge cannot be passed from generation to generation. The probabilities of learning are kept constant and the same for the whole population. Paul G. Higgs studied learning by imitation based on memes [8]. He discusses the action of imitative ability on the process of cultural evolution and finds that genes for increased imitative ability are selectively favored.

For learning, the intelligence of an individual is an important factor. There are some theories about variation of intelligence during life [9,10]. In [11] we represent an evolution model based on bit-string with intelligence. Four genetic strings are introduced and divided into two parts — Bit-String A and Bit-String B. Bit-String A denotes the health of an individual, while Bit-String B describes intelligence. For an individual, the cross reproduction method is used. In that paper, we discuss the population size influenced by intelligence as a function of time steps. In this paper, we study the relation between intelligence and knowledge on the basis of reference [11]. The sexual Penna model with intelligence is introduced at first. After that, we explain how knowledge is represented by another string in our model. The probability of learning will be affected by intelligence. In order to study how the accumulated knowledge influences the survival process by the effect of food and space restrictions, we modify the Verhulst factor. Then, we present the results of our simulations and discuss the evolution of population, intelligence and knowledge respectively. In addition, an equation for calculating the intelligence quotient is given using intelligence and knowledge. We discuss the distribution of intelligence quotient.

[a] e-mail: mfhe@dlut.edu.cn

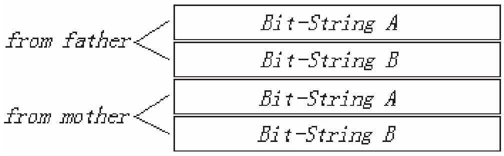

Fig. 1. The individual is represented by four bit strings.

The paper is organized in the following way. In Sections 2 and 3 we develop the Penna model with intelligence and the role of knowledge. Then, we present the results of our simulations and the conclusions are given in the last section.

2 Penna model with intelligence

In this paper we use the sexual Penna model with intelligence [11]. The population is divided into males and females. Each individual is represented by four bit strings of size $A_{\max}$ read in parallel. Two strings are from the father, the others come from the mother. Bit-String A denotes the health of the individual, while the Bit-String B relates to intelligence. The sketch map is shown in Figure 1.

For an individual, genetic diseases are represented by 1 bits in Bit-String A. For Bit-String B the present of a bit "1" represents strong intelligence.

Bit-String A: if an individual has two bits equal to 1 in the same position, it will start to suffer the effects of an inherited disease from that year on until its death. There is a limit number T of diseases each individual can accumulate: if at some age an individual has already acquired T diseases, it dies at that time step. Because the size of strings is $A_{\max}$, each individual will die when its age becomes $A_{\max}$.

Bit-String B: like Bit-String A, if an individual has two bits equal to 1 in the same position, the intelligence is strong. An individual with age i accumulates the sum of strong intelligence bits $B(i)$. This quantity weighs the intelligence level of an individual with age i.

Reproduction: every time step each female with age equal to or greater than the minimum reproductive age R randomly chooses a male with age also equal to or greater than R to mate, generating B offspring. This mating process is repeated every year until death. The offspring genome is constructed in the following way: the father genome is cut in a random position, generating eight bit string pieces. Four complementary pieces, each one coming from one of the original strings, are joined to form the offspring string which contains the genetic charge to be inherited from the mother. After this, M_1 (for Bit-String A) and M_2 (for Bit-String B) random mutations are included. The same procedure is repeated with the mother genome, to produce the other two strings of the baby. The sex of the newborn is randomly chosen. Figure 2 shows this process clearly.

3 The role of knowledge

Now we consider a fifth bit string Bit-String C, also with the same age structure and the same size $A_{\max}$, used to represent knowledge. As opposed to genomic bit strings, this string has all the bits set to 0 at birth, but will change during life. Every time step, an individual has a certain probability P to acquire some knowledge at that age. This knowledge will be represented by a 1 bit in that position of the fifth string; otherwise the 0 bit will be kept, representing no learning at all that year. This string will not be inherited at the process of reproduction. We should expect that the probability of learning would be a function of intelligence. The knowledge relevant in our model is various; not only the basic survival skills but also the stuff we learn later (like theoretical physics). The corresponding knowledge will be learned by individuals with corresponding age. For example, we learn what to eat and how to cross the road when we are very young. They also need to learn the knowledge relevant to their work when they grow up. Thus, we can't say that the learning ability increases as a function of the individual's age. Although the intelligence of an individual will increase as an individual grows, they need to learn more difficult knowledge. So the ability to learn corresponding knowledge for an individual with age i is given by:

$$P = S \cdot \frac{B\,(i)}{i}$$

where $B\,(i)$ is the sum of strong intelligence bits for Bit-String B with age i, and S represents the degree intelligence influences learning.

There are many kinds of knowledge to learn for different kinds of species. For example, the predator will learn catching food and the prey will learn avoiding dangers. In conclusion, knowledge can be used to improve the survival capacity of individuals. As we have seen, in the Penna model an individual dies due to two reasons, accumulation of diseases or the Verhulst factor. The latter represents competition for food or space and depends on a number randomly tossed for each individual at every time step. We consider that the survival rate for an individual will be affected by the knowledge acquired. As an individual grows, they need to learn more and more difficult knowledge and they also need more and more difficult knowledge to continue to survive. So we suppose that all the knowledge needed to learn by individuals with different ages affects the same survival rate for individuals with corresponding ages. Thus, we model the improvement of survival probability for an individual as a function of acquired knowledge $C(i)$ by a new Verhulst factor given by:

$$V = \left(1 - \frac{C(i)}{A_{\max}}\right) \cdot \frac{N_t}{N_{\max}}$$

where $N_{\max}$ is the maximum population size the environment can support and N_t is the current population size. At each time step and for each individual a random number between zero and one is generated and compared with V: if

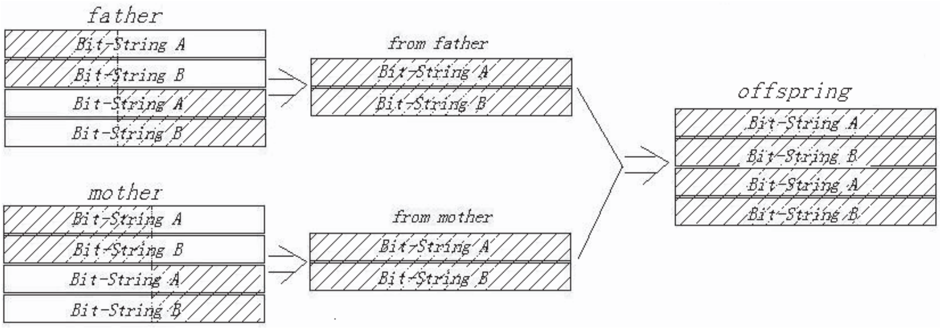

Fig. 2. The construction of offspring genome in the process of sexual reproduction.

Table 1. Values used in the simulations.

Quantity	Value
Maximum population size	$N_{\max} = 10^5$
Initial population size	$N_0 = 1000$
Size of the bit strings	$A_{\max} = 32$
Maximum number of deleterious mutations	$T = 4$
Minimum reproduction age	$R = 8$
Mutation rate of Bit-String A	$M_1 = 2$
Mutation rate of Bit-String B	$M_2 = 2$
Birth rate	$B = 1$

it is smaller than V, the individual dies independently of its age or genome.

All these processes of testing the survival of each individual and the process of reproduction, both applied over the whole population, represents a time step, i.e., one year in the simulation.

4 Results

The values used for all the simulations shown in this paper are given in Table 1.

We discuss the evolution of population, intelligence and knowledge respectively in the following.

4.1 The evolution of population

The evolution of population is shown in Figure 3 with different value of S.

From Figure 3 we can see that after many time steps stability is reached and the population self-organizes. When the value of S is small, the equilibrium will almost be the same for different S. But if S reaches 1, the equilibrium will increase obviously.

The symbol $A_w^k(i)$ expresses that the value of the ith position of Bit-String A for individual$k$. When $w = m$, the string that comes from the father will be considered. If $w = f$, we will make use of the other string from mother. Where $A_w^k(i) = 0$ or $A_w^k(i) = 1$. The average status of

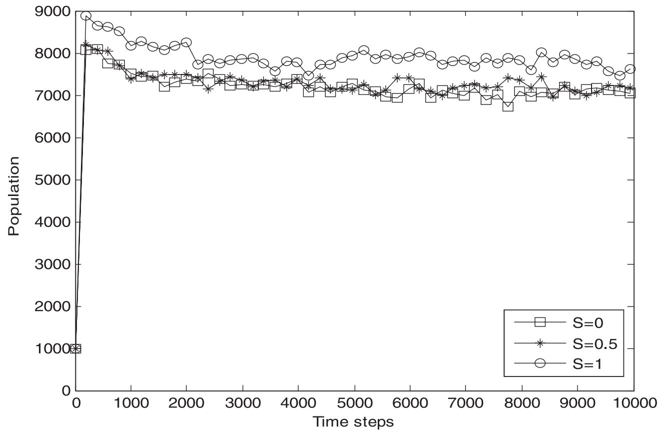

Fig. 3. Population size as a function of time step.

disease for an individual at t time step can be given by:

$$\alpha_t = \frac{\sum_{k=1}^{N_t} \sum_{i=1}^{a(k,t)} sign\left(A_f^k(i) \cdot A_m^k(i)\right)}{N_t},$$

where

$$sign(x) = \begin{cases} 0 & x = 0 \\ 1 & x \neq 0 \end{cases},$$

$a(k,t)$ denotes the age of individual$k$ at t time step and N_t means the population size at t time step. Figure 4 shows the curve of α_t.

From Figure 4 we can see the average status of disease will decrease with increasing S. It means that learning will make individuals healthier against disease.

After many time steps, stability is reached and the population self-organizes. Stability means that the average number of individuals of any given age is constant at each time step. So the survival rate with age a can be given by $S(a) = N_{t+1}(a+1)/N_t(a)$ where $N_t(a)$ is the number of individuals with age a at t time step, The survival rate represents the probability an individual with age a has to reach age $a+1$. This curve of survival rate will be shown in Figure 5.

From Figure 5 we can see that the survival rate of individuals will decrease slowly at smaller ages. But with increasing age, it decreases more rapidly. In addition, when the value of S is small, it is not obviously that learning

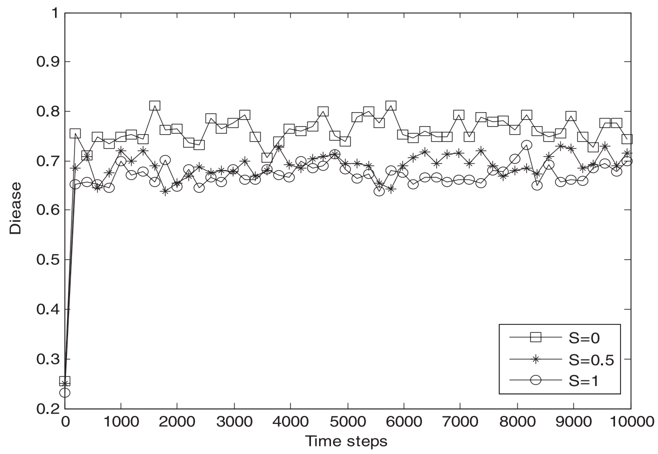

Fig. 4. The average status of disease as a function of time.

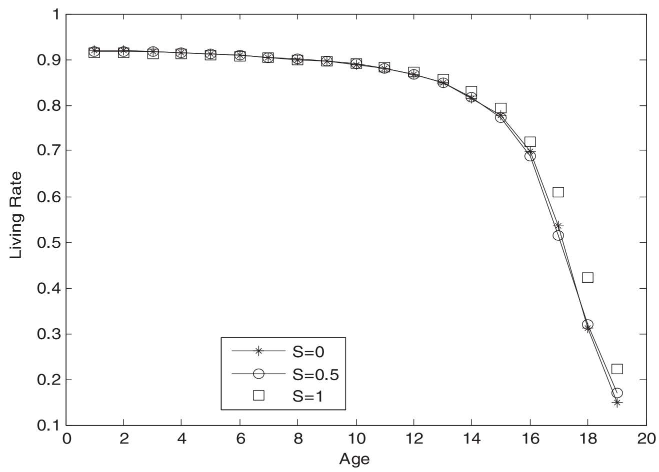

Fig. 5. The survival rate of individuals as a function of age.

knowledge can increase the survival rate. But on the condition of $S = 1$, the survival rate will enhance for older individuals. That means the intensity of learning contributes to survival rate.

The symbol $\beta_t(i)$ denotes the proportion of genes with disease for homozygote in position i at t time step. So $\beta_t(i)$ can be described by:

$$\beta_t(i) = \frac{\sum_k^{N_t} sign\left(A_m^k(i) \cdot A_f^k(i)\right)}{N_t}.$$

After many time steps, stability is reached and the population self-organizes. When the system steadies, $\beta_t(i)$ at each time step will be the same. It means that $\beta_t(i) = \beta(i)$. The distribution of $\beta(i)$ will be shown in Figure 6.

From Figure 6 we can see that the proportion is very low in the position before R. That will be as a result of natural selection. After the position R, the proportion increases quickly and reaches a hundred percent swiftly. That will be as a result of accumulation. Furthermore, the trend of increasing is almost similar with different S. But the position of increasing will be delayed a few bits as the

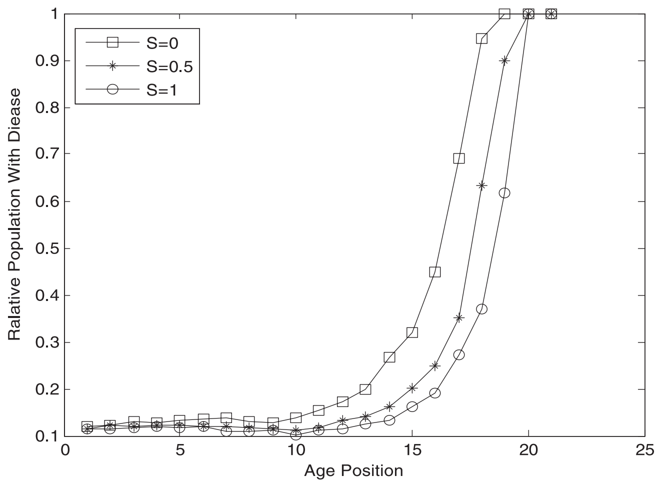

Fig. 6. The proportion of genes with disease for homozygote as a function of age position.

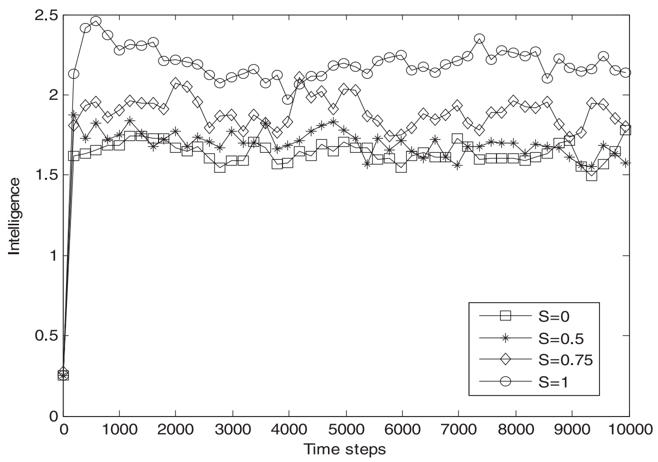

Fig. 7. The average level of intelligence for an individual as the function of time.

increasing of S. That will be as a result of the learning effect.

4.2 The evolution of intelligence

The symbol $B_w^k(i)$ expresses the value of the ith position of Bit-String B for an individual k. When $w = m$, the string that comes from the father will be considered. If $w = f$, we will make use of the other string from the mother. Where $B_w^k(i) = 0$ or $B_w^k(i) = 1$. The average level of intelligence for an individual at t time step can be given by:

$$\lambda_t = \frac{\sum_{k=1}^{N_t} \sum_{i=1}^{a(k,t)} sign\left(B_f^k(i) \cdot B_m^k(i)\right)}{N_t}.$$

Figure 7 shows the curve of λ_t.

From Figure 7 we can see that after many time steps stability is reached. It means that the average level of intelligence can not change obviously over time. Nevertheless, as S increases, the value of steady state will increase

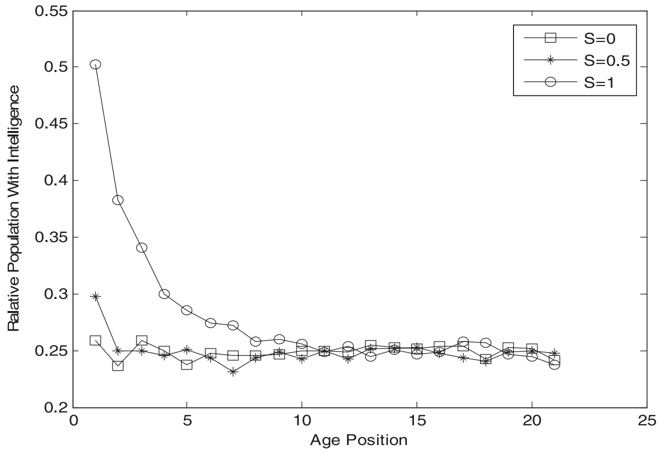

Fig. 8. The proportion of genes with intelligence for homozygote as a function of age position.

gradually. It can be concluded that learning knowledge will contribute to the evolution of intelligence. Though there is no direct relation between S and the intelligence of individual.

The symbol $\delta_t(i)$ denotes the proportion of genes with intelligence for homozygote in position i at t time step. So $\delta_t(i)$ can be described by:

$$\delta_t(i) = \frac{\sum_{k}^{N_t} sign\left(B_m^k(i) \cdot B_f^k(i)\right)}{N_t}.$$

When the system steadies, $\delta_t(i)$ at each time step will be the same. It means that $\delta_t(i) = \delta(i)$. The distribution of $\delta(i)$ will be shown in Figure 8.

The proportion of genes with intelligence for homozygote decreases gradually at an early age. At a high age stability will be reached and the fraction of homozygote "11" is 0.25, which means random sequences. In addition, at an early age, the larger value of S, the larger proportion of genes with intelligence. It shows that for children the increasing of learning ability can enhance the average level of intelligence. But for adults, there is not any visible effect. This relates to what is known about early learning leading to development of intelligence. The bits for the early years contribute to intelligence in all the later years. Therefore there is stronger selection acting on the early bits. It also shows the importance of early learning.

4.3 The evolution of knowledge

The symbol $C_k^t(i)$ denotes the value of the ith position of Bit-String C for individual k. Where $C_k = 0$ or $C_k = 1$. The average level of knowledge for a species can be described by:

$$\gamma_t = \frac{\sum_{k=1}^{N_t} \sum_{i=1}^{a(k,t)} C_k(i)}{N_t}.$$

The distribution of $\gamma_t$is shown in Figure 9.

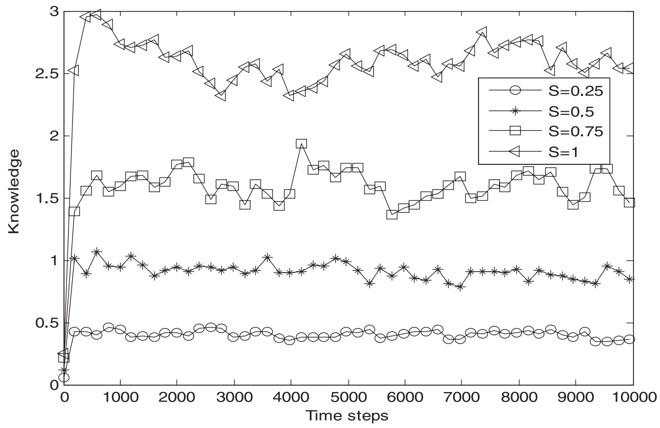

Fig. 9. The average level of knowledge for a species as the function of time.

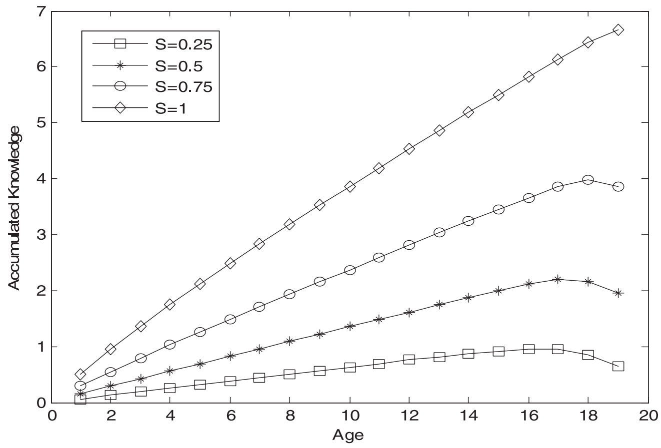

Fig. 10. Knowledge accumulated as a function of age.

From Figure 9 we can see that the average level of knowledge for a species tends to stabilize. The value of knowledge at steady state will increase with increasing S. It means that increasing the learning ability will enhance the average level of knowledge for a species. But there is no contribution to the average level of knowledge with time.

The symbol $C^t(a)$ expresses the knowledge accumulated with age a at t time step. The $C^t(a)$ can be given by:

$$C^t(a) = \frac{\sum_{k=1}^{N(a)} \sum_{i=1}^{a} C_k^t(i)}{N(a)}.$$

When stability is reached, $C^t(a)$ at each time step will be the same. It means that $C(a) = C^t(a)$. The distribution of $C^t(a)$ will be shown in Figure 10.

Figure 10 shows that knowledge accumulated as a function of age. They fit basically a linear function. In addition, the slope of line increases with increasing S. It means that though there are no distinct differences for the average level of knowledge for a species as a function of time (Fig. 9), the knowledge accumulated increases as a function of age. But because of the high death rate with elders, the curve will decrease appreciably.

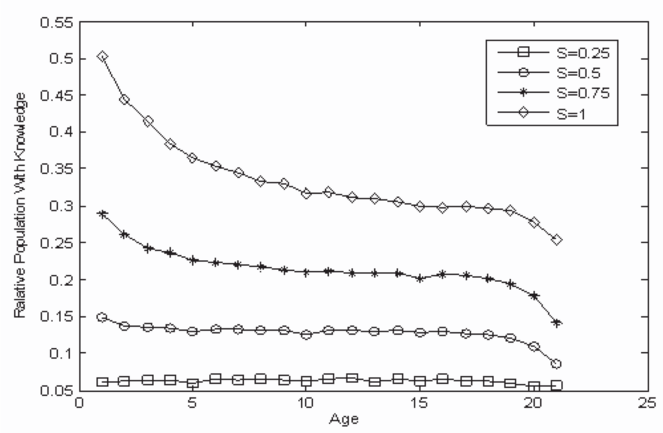

Fig. 11. The proportion of bits with knowledge as a function of age position.

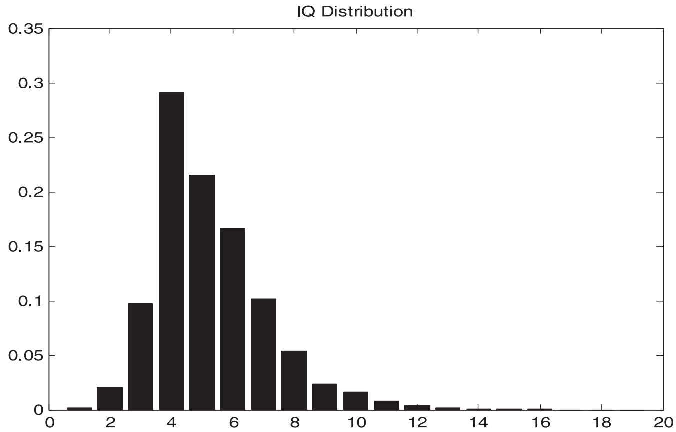

Fig. 12. The distribution of intelligence quotient, where $S = 0.5$ and $\alpha = 0.7$.

The symbol $\eta_t(i)$ shows the proportion of bits with knowledge in position i at t time step. So $\eta_t(i)$ can be described by:

$$\eta_t(i) = \frac{\sum_{k=1}^{N_t} C_k^t(i)}{N_t}.$$

When the system steadies, $\eta_t(i)$ at each time step will be the same. It means that $\eta_t(i) = \eta(i)$. The distribution of $\eta(i)$ will be shown in Figure 11.

Figure 11 shows that the proportion of bits with knowledge as a function of age position. From this figure we can see that the proportion decreases in the lower age range. The reason is that there are more possibilities to learn knowledge for youth. After some years, stability is reached and lasts for a long time. At this time the quantity of knowledge acquired will not reduce though it does not increase. It means that the individual will acquire knowledge steadily in this period. The figure also shows that the quantity of knowledge acquired increases with increasing S. That will be as a result of the learning effect. But if the value of S reaches 1, the quantity of knowledge acquired will reduce with increasing age.

4.4 The intelligence quotient

The individuals we introduced have the character of intelligence. The rules of learning are set up in our paper. So we can weigh the ability of individuals by a certain index (intelligence quotient). Parts of intelligence quotient will come from inheritance and the others will be affected by environment [12]. We use the information of Bit-String B (intelligence) to describe the parts from inheritance. The other parts from the environment will be expressed by the Bit-String C (knowledge). The intelligence quotient IQ can be given by:

$$IQ = \alpha \frac{B(A_{\max})}{A_{\max}} + (1-\alpha)\frac{C(i)}{i},$$

where α is the percentage of the parts from inheritance. The distribution of intelligence quotient is shown in Figure 12, where $\alpha = 0.7$.

From Figure 12 we can see that the proportion of population with low intelligence quotient is small. Moreover, there are more individuals with high intelligence quotient. But the increasing speed is fast for low intelligence quotient. Though the distribution is not precisely a normal distribution, the basic character of normal distribution can be seen in this figure.

5 Conclusions

An evolutionary model based on Bit-String with intelligence and learning is set up in this paper. Each individual is represented by five bit strings. Four of them relate to genes and the other denotes the knowledge from learning. The four strings relative to genes are divided into two parts — Bit-String A and Bit-String B. Bit-String A denotes the health of an individual, while Bit-String B describes intelligence. For an individual, the cross reproduction method is used in this model. After that we explain how knowledge is represented in our model. The probability of learning will be affected by intelligence. In order to study how the accumulated knowledge influences the survival process by the effect of food and space restrictions, we modify the Verhulst factor. Then, we present the results of our simulations and discuss the evolution of population, intelligence and knowledge respectively. In addition, an equation to calculate the intelligence quotient is given out base on intelligence and knowledge. And we discuss the distribution of intelligence quotient.

Some conclusions can be obtained from this model.

1. After many time steps, stability is reached and the population self-organizes.
2. Learning knowledge makes individuals healthier against disease and enhances the living rate to a certain extent.

3. After many time steps, the average level of intelligence also tends to stability. Learning knowledge contributes to the evolution of intelligence, especially for children.
4. The increasing of learning ability will enhance the average level of knowledge for a species. But there is no contribution to the average level of knowledge as a function of time.

We are grateful to reviewers for their helpful comments.

References

1. E. Gonzalez et al., Science **307**, 1434 (2005)
2. Hannes Lohi et al., Science **307**, 81 (2005)
3. Virginie S. Vervoort et al., Science **296**, 2401 (2002)
4. T.J.P. Penna, J. Stat. Phys. **78**, 1629 (1995)
5. S. Moss de Oliveira, P.M.C. de Oliveira, D. Stauffer, Braz. J. Phys. **26**, 626 (1996)
6. D. Stauffer, P.M.C. de Oliveira, S. Moss de Oliveira, T.J.P. Penna, J.S.S. Martins, An. Acad. Bras. Cienc. **73**, 15 (2001)
7. A.T. Bustillos, P.M. de Oliveira, Phys. Rev. E **69**, 021903 (2004)
8. P.G. Higgs, Proc. Roy. Soc. London Ser. B **267**, 1355 (2000)
9. C.T. Morgan, *A Brief Introduction to Psychology* (McGraw-Hill, New York, 1974)
10. H. Kaplan, A. Robson, Proc. Natl. Acad. Sci. USA **99**, 10221 (2002)
11. M. He, Q. Pan, B. Yu, Int. J. Mod. Phys. C **16**, 1303 (2005)
12. R. Grant Steen, *DNA and Destiny: Nature & Nurture in Human Behavior* (Plenum Publishing Corporation, 1996)