Automatic Generation of Children's Songs Based on Machine Statistic Learning

Li Pan
Mudanjiang Normal College, Mudanjiang, China
`panli_1982@163.com`

Abstract—In this paper, the automatic generation of children's songs is studied. First of all, according to the number of key lyrics submitted by the songwriter, statistical machine learning is used to expand and obtain more theme-related lyrics, and then the first sentence is generated through the language model automatically. On this basis, the subsequent sentences are generated through the statistical machine learning translation method. In the process of the generation, the statistical machine learning is used to expand the conception of the song, so as to get richer sentence candidates. The main features and contributions of the study are: Firstly, the statistical machine learning translation is put forward as the theoretical basis, the preceding and next sentence relationship of children's songs are mapped into the relation of the source language and target language in the statistical translation model, and the machine statistics learning translation model is designed with the integration of the domain knowledge of songs. Secondly, the statistical machine learning is used in the generation process to expand the lyrics words, thereby enhancing the theme and conception of the song. The experimental results have confirmed the effectiveness of the proposed method.

Keywords—Generation of Songs; Statistical machine Learning; Statistical machine Learning Translation; Automatic Evaluation

1 Introduction

Children's songs are a kind of songs, and the creation of children's songs focuses on beautiful words and metrical rhyme, requires completing the description of the theme content within the specified word count, and contains the specific conception and emotional expression. However, the wording and phrasing of children's songs, rhythm and other constraints are the major obstacle in the learning and creative process for the literary lovers without special training. We carry out the statistical machine learning of the large scale song data, integrate the knowledge of children's songs into the statistical probability model, and make use of the computer to assist in the creation of children's songs, which has not only provided support for the massive music culture lovers, but also has positive significance for the transmission and pub-

licity of the Chinese culture. In addition, as the research in the field of natural language generation, this paper has also provided new ideas for reference.

In this paper, the automatic generation of children's songs is studied. First of all, according to a number of theme lyrics submitted by the songwriter, statistical machine learning is used to expand and obtain more theme related lyrics, and then the first sentence is automatically generated through the language model. On this basis, the statistical machine learning translation method is used to generate the subsequent sentences. In the process of generation, the statistical machine learning is applied to expand the conception of the songs, so as to get richer sentence candidates. The main features and contributions of this study are: Firstly, the statistical machine learning translation is put forward as the theoretical basis, and the preceding and next sentence relationship of the children's song is mapped into the relationship between the source language and the target language in the statistical translation model. Taking the melody and rhyme into consideration, the statistical machine learning translation model is designed to incorporate the domain knowledge of songs. Secondly, the statistical machine learning is used to expand the lyrics words in the generation process, thereby enhancing the theme and conception of the song. And the experimental results have confirmed the effectiveness of the proposed method.

2 Related Work

Foreign song generation research started in 1959, when Lutz [2] used the computer to generate the first song in German. The methods of song generation can be classified into the generation method based on template, generation and testing method, method based on the genetic algorithm, and the reasoning method based on instance:

1. Based on the template generation method, a template is given for the composing of songs when the rhythm syntax and other constraints are met.
2. Generate and test the methods. Generate the random lyrics sequence according to the formal requirements, and use the corresponding constraints and evaluation criteria to determine whether the sequence meets the requirements. Manurung's Chart system [3], WASP system [4] and Tra- la - Lycics [5] are all based on this approach.
3. Method based on genetic algorithm. Combining the genetic algorithm and the evaluation module, the generation module generates the candidate works using the genetic algorithm according to the syntax and other information, and the evaluation module scores the candidate output according to the certain criteria. The representative systems include POEVOLVE [6] and McGonna g all [7].
4. Method based on instance reasoning. Through retrieving the existing songs, the content of the existing songs is adjusted according to the target information to be described by the user. ASPERA [8] and COLIBRI [9] are representatives of such method systems.

Domestic song generation research started in the mid 90s of the 20th century, and has accumulated a lot of predecessors' work so far. For example, Taiwan Luo Feng-

zhu's metrical check and rhyme lyrics search system. In the studies of Zhou Changle et al. [10] in the song generation, the method was to perform automatic song generation using the genetic algorithm on the basis of the given lyrics plate and prosodic template. In the aspect of Chinese couplet generation, the computer automatic couplet system developed by Microsoft Asia Research Institute of Natural Language Computing Group [11] applied the statistical machine learning translation to the automatic generation of the second line of the couplet. He et al. [12] drew on the idea of the generation of the second line from the first line of the couplet, and extended it to the automatic generation of children's songs. Genzel et al. [13] once also applied the statistical machine learning translation idea to the machine translation of songs with rhythmic constraints.

In this paper, the method of statistical machine learning translation for the generation of children's songs is adopted. The preceding and next sentence relationship of the children's songs is regarded as the correspondence relationship between the source sentences and target sentences in statistical machine learning translation, and reasonable specialization is conducted to the statistical machine learning translation model, so that it is consistent with the characteristics of the generation of the children's songs. By learning the creation knowledge of the songwriting from the training corpus, the automatic generation of the next sentence can be realized in the case of the existence of the preceding sentence. The children's songs generation method based on the statistical machine learning translation of He et al. [12] requires the use of an artificial semantic classification dictionary - "Theories of lyrics in songs", the creator must select the key words in the dictionary to generate the first sentence, which is easy to result in the constraints to the creation themes. And as "Theories of lyrics in songs" was written in the 17th century, many of the lyrics are no longer in common use, which are easy to generate songs hard to understand. In addition, the statistical machine learning process does not take the consistency of the generated content and the presentation of the theme into consideration, therefore, the relevance of the generated songs and the themes is not taken into account as a feature. The result is rhetorical; however, it is often very difficult for the readers to appreciate the relatively consistent theme or conception. Finally, due to the lack of automatic evaluation method, the logarithmic linear model used for the decoding in the statistical machine learning translation does not make parameter adjustment yet, but based on the parameters set by the experience, this method may not be able to search the local optimal solution of the model.

This paper is an improvement on the work of He et al. [12]. And the main features and contributions are as the following: (1) The statistical machine learning translation is put forward as the theoretical foundation, the relationship between the preceding and next sentence of children's songs is mapped to the source language and target language in the statistical translation model, and the statistical machine learning translation model is designed to integrate the domain knowledge of songs; (2) Introduction of the statistical machine learning as the semantic related extension of the related lyrics set, and the songwriter should be able to enter the key lyrics freely, making the creation of the themes more freely and extensively. Statistical machine learning is used to measure the semantic correlation between the generated songs and the theme

lyrics that is generate, which has enhanced the coherence of the content of the songs generated and the theme expression, and thus strengthened its artistic conception.

3 Effect of Machine Statistics Learning on the Assisted Generation of Children Songs

This section introduces the PLSA and its training method, and expounds the expansion method of the semantic related lyrics sets, as well as the semantic relevance measure between the generated songs and theme lyrics, at the same time, gives the algorithm for the generation of the first sentence according to the theme lyrics.

3.1 Creation Architecture of Children's Songs

Generally, the creation process of the children's songs can be broadly divided into the following steps:

1. Have clear description of the themes, and these themes can be expressed through the specific theme lyrics. For example, writing a hometown missing song, the words such as "Homesickness" or "Hometown" and other specific words can be used to express the theme.
2. According to the theme lyrics more semantically related lyrics are associated, such as from "Homesickness" associating to "Beloved mother" and "A letter from home" and so on.
3. When the constraints for the children's songs are met, consider the creation of songs through the reasonable organization of these lyrics.
4. In the process of completion of the whole song, refer to the preceding lyrics and theme to create the next sentence step by step, run in cycles until the whole song is completed.

Accordingly, we design the specific structure of the children's songs assisted creation system as the following (Figure 1):
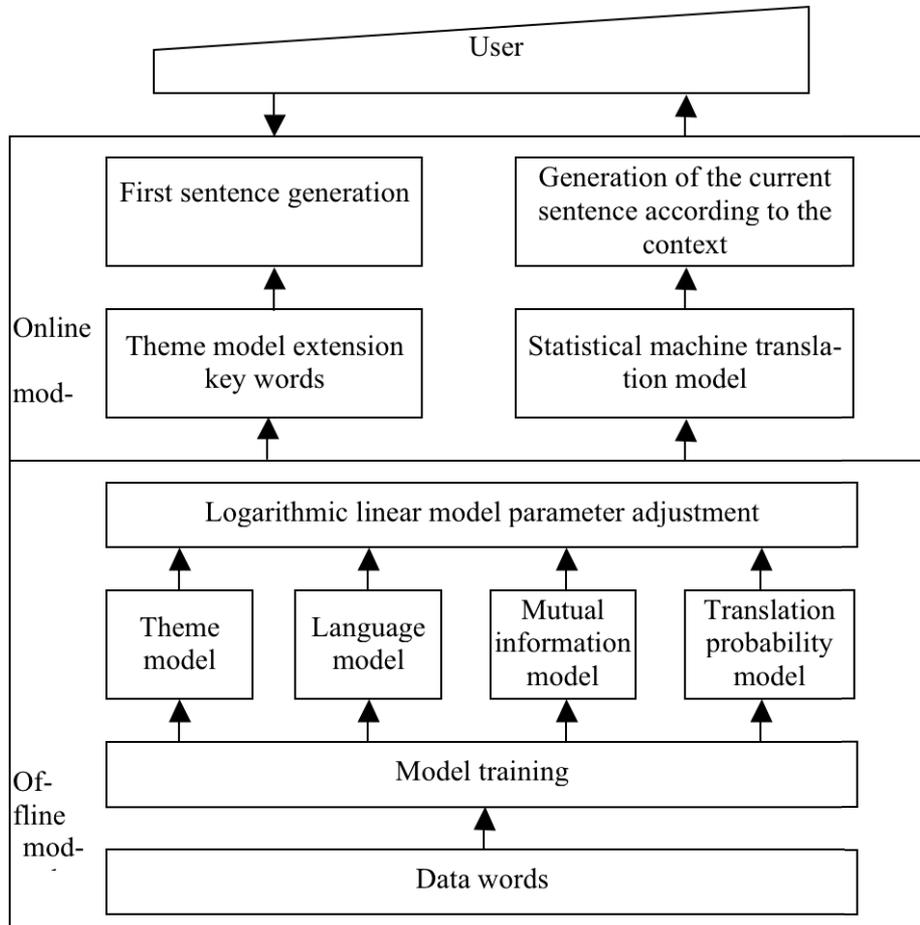
Fig. 1. Children's songs assisted creation system architecture

3.2 PLSA and its Training

PLSA is an automatic document indexing method based on the statistical machine statistics learning information. It can deal with the problems of lyrics clustering and document clustering for the domain oriented synonymous lyrics or polysemy lyrics. Different from the co-occurrence relationship between the document and the lyrics obtained from the training corpus, PLSA makes use of the potential theme of the unknown variable, through analyzing the co-occurrence relationship among the potential theme, document and lyrics, to indirectly construct the co-occurrence relationship between the document and lyrics. The theory has two independent hypotheses: the document and the lyrics in the observation data do not have the bag-of-words; and the lyrics generated after the given theme are irrelevant to the document. From the perspective of the model generation, PLSA can be described as the following:

1. Select the document d from the document set with the probability $P(d)$;

2. The document describes the theme t with probability $P(d|t)$;

3. Describe the theme t using with the lyrics w with the probability $P(w|t)$.

Then the relationship between the lyrics and the final document can be established through a joint probability formula, as the following:

$$
\begin{aligned}
P(d,w) &= P(d)P(w|d) \\
&= P(d)\sum_{t\in T} P(w|t)P(t|d) \\
&= \sum_{t\in T} P(w|t)P(d|t)P(t)
\end{aligned}
\tag{1}
$$

The idea of maximum likelihood is used to make estimation for $P(w|t)$, $P(d|t)$ and $P(t)$. The maximum likelihood estimation process can be implemented by the EM (Expectation Maximization) algorithm. Finally, the co-occurrence matrix between the training set theme and the lyrics as well as the co-occurrence matrix between the theme and the document are obtained. Pre initialization $P(t)=\dfrac{1}{K}$, $P(w|t)=\dfrac{1}{N}$ and $P(d|t)=\dfrac{1}{M}$ (Where K is the predefined number of potential themes, M is the number of the documents contained in the training data, and N is the number of lyrics), and then EM iteration is used for the optimization of the estimated values. EM iteration steps: Calculates the probability distributions of the observed lyrics and the document on each theme according to the current estimation parameters; further optimize the estimation of the model parameters by the probability distributions of the lyrics and the document on the theme. And the EM iteration equation is as the following:

E-Step:

$$
P(t|d,w) = \frac{\left(P(t)P(d|t)P(w|t)\right)^{\beta}}{\sum_{t}\left(P(t')P(d|t')P(w|t')\right)^{\beta}}
\tag{2}
$$

M-Step:

$$
P(w|t) = \frac{\sum_{d} n(d,w)P(t|d,w)}{\sum_{d,w'} n(d,w')P(t|d,w')}
\tag{3}
$$

$$P(d|t) = \frac{\sum_{w} n(d,w) P(t|d,w)}{\sum_{d',w} n(d',w) P(t|d',w)} \tag{4}$$

$$P(t) = \frac{1}{R} \sum_{d,w} n(d,w) P(t|d,w) \tag{5}$$

In which $R = \sum_{d,w} n(d,w)$, $n(d,w)$ refers to the number of times that the lyrics w occurs in the document d. In this paper, TEM algorithm (Tempered EM) is adopted, compared with the conventional EM algorithm, by adding the coefficient β to the estimation equation (2), it can effectively prevent the over-fitting of the parameters. After the training is completed, the document-theme co-occurrence matrix and the lyrics-theme co-occurrence matrix can be obtained. In the co-occurrence matrix, each document or lyrics can be described by a theme distribution vector whose dimension has the potential theme number of K. The corresponding theme distribution vector of lyrics or documents is used, combined with the similarity measurement method, to perform lyrics, document clustering or other operations in the vector space model. For example: There are six documents d1~d6 and two potential themes t1 and t2. After PLSA training is completed, the potential theme space of the document is represented in Figure 2. If the semantic relevance is calculated by Euclidean distance, the semantics relevance of d1 and d2 is greater than the semantic relevance of d1 and d6. It can be deduced from the figure that the theme contents described by d1, d2 and d3 are relatively relevant, while compared with the theme contents described by d4, d5 and d6, there will be relatively huge difference.
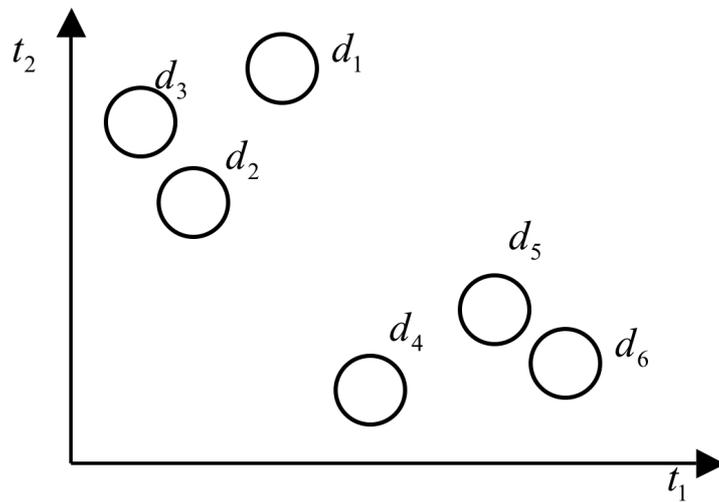

Fig. 2. Example of the document in the two-dimensional latent semantic space

3.3 Lyrics Words Extension Based on the Statistical machine Learning

In children's song writing, a variety of lyrics words can be used to describe a given theme. In order to get rich descriptive lyrics set of the target songs, we make the extension to the lyrics of the key songs through the selection of the semantic related lyrics of the songs (semantic relevance refers to the common lyrics of lyrics A and lyrics B that belong to the same theme, which are often mixed in the artificially composed article of a certain class of theme, and there are a large number of co-occurrence relationships in the training document library). After the experimental comparison, compared with the general lyrics clustering method based on lyrics context, the PLSA-based clustering method is more robust. By mapping the relationship between high dimensional lyrics to the low dimensional lyrics and theme relationship, the lyrics that are original lack of co-occurrence information and thus not established with the semantic relevance can also be analyzed and represented by the PLSA and hence establish the semantic relationship; moreover, PLSA training process presupposes that the number of potential themes is equal to the value of the size of the dimension after the lyrics in the co-occurrence matrix are represented by the potential theme, and its value taking affects the lyrics clustering results. By adjusting the number K of the potential themes, the strength of the semantic relevance between the lyrics can be controlled. The higher the latent theme distribution vector dimension of the lyrics is, the more scattered they are in the spatial distribution, and the stricter the corresponding concept of semantic relevance shall be. The lower the dimension is, the denser the spatial distribution is, and the rougher the concept of semantic correlation will be.

In this paper, the co-occurrence matrix of lyrics and theme is obtained by using PLSA, and then the semantic relevance between the lyrics is measured according to the theme distribution vector that the lyrics correspond to. Table 1 shows the semantic related lyrics of the given lyrics entry that different Ks correspond to, and K is the number of potential themes, from which it can be seen that, with the increase of K, the expansion of the lyrics will be more closely related to the lyrics of the given theme. Considering from the demand of the expression diversity and the calculation of the cost balance, in this paper, K is taken as 50.

Table 1. Examples of semantic related lyrics

K\Lyrics entry	Autumn Water	Distant Hills	Running Stream
K=10	Sideburn, eulogistic, grand sound, surge, zither music, temple, officer, omen, cavity	Long journey, implicit, bird nest, to each, blocking, belt, bake, store, conceal	Central plain, fortune, comfortably drunk, serenade, entry, brow, boat, regretful, yet
K=50	Amazing shadow, lapel, arbor, lonely image, late autumn, desolate, lock, pursue, venue	Mountains, exhibition hall, autumnal scenery, late crow, smoke flow, maple, riverside, float, green jade	Song-dance drama, lucid brook, bright moon, first floor, instruction, tie, follow, seek, go
K=100	Lean against the railing, shine, jade flute, moon glory, flute music, quiet, cold, broken, skinny	Light cloud, a few pieces, trail, eau sauvage, after rain, cage, chirp, loom, fluffy	Duckweeds, expel water, passionately attached, purple street, falling site, pass, seek, follow, likelihood

Theme lyrics extension method. For each of the given theme lyrics, the lyrics with the semantic relevance greater than 0.6 are selected from the semantic relevance lyrics database and added into the extended lyrics set (the range of relevance is [0, 1]). For example, given two theme lyrics A and B, remove the duplication of A corresponding semantically related lyrics and B related semantics relevant lyrics, and then take the union and add into the extended lyrics words set.

3.4 Semantic Relevance Measurement Method for the Generated Songs and Theme Lyrics

In order to measure the semantic relevancy between the newly generated songs and the lyrics of a given theme, this paper calculates the semantic relevance of the two by referring to the principle of statistical machine learning application to the information retrieval. The method is as the following: The theme lyrics is regarded as a special short document, which is mapped into potential theme space by EM iterative process in the PLSA model training process, and the corresponding theme distribution vector is obtained and denoted as $P(t|q)$. Carry out the same operation for the newly generated song and denote it as $P(t|d)$, the semantic relevance between the theme lyrics and the generated songs can be calculated by the cosine similarity.

The method of mapping the new document to the potential theme space is as the following: In the EM algorithm in Section 4.1, $P(t)$ and $P(w|t)$ are fixed and $P(t|d,w)$ and $P(d|t)$ in Equation (2) and (4) are iterated and updated until the convergence or the iteration times are up. Table 2 shows the semantic related songs retrieved based on the lyrics "Campaign".

Through to the semantic similarity calculation, the song content retrieved according to the key lyrics semantics shall be closely related to the theme and has the corresponding conception. By reverse thinking, if in the process of generation, the semantic relevance measurement on the candidate songs is strengthened, the content thus generated can be more closely related to the theme and reflect the conception. To this end, in the statistical machine learning translation model, the feature of theme relevance is added in this paper.

4 Statement Generation Model Based on Statistical Machine Learning Translation

In recent years, the studies of statistical machine learning translation have developed very rapidly, which are attributing to the increase of the training data and the emergence of different translation models. In general, the current mainstream statistical machine learning methods can be divided into the following types of models: String-to-string, tree-to-string, string-to-tree and tree-to-tree. Different models have their own characteristics, but also have their own strengths and weaknesses. Jiang et al. once applied the statistical machine learning translation algorithm to the generation

of the second line of the Chinese couplet [11]. He et al. [12] once applied the statistical machine learning translation to the generation of children's songs. In this paper, after the first sentence is generated, the second, third and fourth sentence are generated based on the statistical machine learning translation method; and the new features are added to achieve the expected effect of consistent lyrics and conception of the entire song.

Phrase-based Statistical Machine Translation (hereinafter referred to as PBSMT for short) is one of the most mainstream machine translation techniques. Its advantage lies in the accuracy of the selected lyrics in the phrase translation results. As the generation of songs emphasizes the antithesis, which does not involve the distant word order adjustment issue, therefore, the generation of songs is very suitable to use the phrase-based machine translation algorithm to solve. At present, the mainstream statistical machine learning algorithms are all based on the maximum entropy framework, that is, given the source language sentence f, by calculating and perform statistics on the different translation characteristics, using the maximum entropy model to calculate the probability of the generated target language sentence e. Then sort by the probability, and select the translation candidate with the maximum probability as the result, the formal description is as the following:

$$
\begin{aligned}
\hat{e} &= \arg\max_{e} \left\{ P(e|f) \right\} \\
&= \arg\max_{e} \left\{ \frac{\exp\left[\sum_{m=1}^{M} \lambda_m h_m(e,f) \right]}{\sum_{e} \exp\left[\sum_{m=1}^{M} \lambda_m h_m(e',f) \right]} \right\} \\
&= \arg\max_{e} \left\{ \sum_{m=1}^{M} \lambda_m h_m(e,f) \right\}
\end{aligned}
\tag{6}
$$

In which, $h_m(\cdot)$ stands for the characteristic function in the maximum entropy model, λ_m stands for the corresponding feature weight, and M stands for the number of feature functions. As only the translation candidates with the highest probability are expected to be obtained, the specific probability value is not concerned, thus the denominator portion in the above equation is removed, and the equation still holds.

In the statistical machine learning translation, although the choice of model is important, the choice of feature function cannot be ignored, because it directly determines whether the model can generate the correct translation results. Generally speaking, the phrase-based statistical machine learning translation system uses the following feature functions: Forward phrase translation probability, reverse phrase translation probability, positive phrase lexicalization translation probability, reverse phrase lexicalization translation probability, number of phrases used and target language model. In addition to the aforementioned basic features, this paper also introduces two new features: Mutual information features and machine statistics learning for the lyrics words and song related features, in which, the mutual information features have been used in the Chinese couplet generated, and the experimental results have shown

its effectiveness; The statistical machine learning feature is used in the field of song generation for the first time, as shown in Table 2.

Table 2. Feature function of the song generation model

Feature Function	Description
$h_1(e,f) = \prod_{\forall i,j;s.t.;i+1=j} MI(e_i, e_j)$	Mutual information of two continuous lyrics in the generated songs
$h_2(e,f) = \prod_{i=1}^{l} P_{topic}(e_i \mid t)$	Relevance of the generated songs and the given theme

The two new features can further help the song generation model optimize the lyrics, and expect that different songs can achieve better consistency of theme conception. $P_{topic}(e_i \mid t)$ stands for the semantic correlation between the lyrics ei and the theme t, and t is the theme distribution vector that projects one or more theme lyrics as short documents to the theme space, where l is the number of words in the sentence.

Firstly, the sub lyrics strategy is used to randomly initialize the lyrics and lyrics frequency information, and the lyrics frequency is used to segment the children's songs corpus; then, the estimation of the information of the lyrics frequency information is optimized on the corpus after segmentation, and the lyrics with the lyrics frequency lower than the set threshold are filtered (take 1.0E-8 in this paper); repeated iteration is carried out until the lyrics frequency converges, and the lyrics list after the convergence and the lyrics frequency information is used as the background sub-lyrics model. Considering the characteristics of the children's songs that the maximum length of a single sentence is 7, this paper adopts the sub-lyric strategy based on the unary grammatical probability. The probability value of all kinds of segmentation of the input sequence is calculated by the sub-lyrics model, and the one with the maximum probability is selected as the final segmentation sequence.

The decoding algorithm refers to the method of Koehn et al. [15]. At the same time, in order to meet the constraint of rhyme, in the fourth sentence of the decoding, according to the last word of the second sentence, delete the candidate lyrics that do not meet the prosodic constraint. In addition, to ensure that there are enough candidate lyrics, this paper adds the lyrics that meet the prosodic constraints. The decoding algorithm introduces the above features into the logarithmic linear model. After selecting the performance evaluation method, the minimum error rate training algorithm [14] is used to train the parameters. Then the model uses the trained parameters to generate the next sentence for the given song. The specific parameters training details for the logarithmic linear model will be introduced in the next section. It should be noted that, this paper constructs the model for the preceding and next sentences in different positions of a song, namely, training the translation model of the first and second, second and third, third and fourth sentence respectively.

5　　Experiment

This section describes the training data, experimental design, and results analysis for each statistical model. The feasibility of the replacement of the "Theory of Lyrics of Songs" with the semantic relevance lyrics database generated by the statistical machine learning is validated, and the results of the manual evaluation and automatic evaluation of the children's songs are analyzed. Finally, several samples of children's songs generated by the system are illustrated.

5.1　Training Data of the Statistical Model

Children's song training corpus come from the Internet, including "Children's Songs", "Taiwan Songs", "All Taiwan Songs" and "Mainland Songs Selection" and other literatures, as well as the children's songs after extraction and screening from all the major song forums (such as songs online, Tianya forum song improvising), totally more than 287,000 songs. The first and second, second and third, third and fourth sentences of each song in the corpus train the corresponding translation models respectively, including the positive and negative phrases and the lexicalized translation models. All the single sentence songs are used to train the language models and the mutual information models in the single sentence. The different sentences in a song are used to train the mutual information model between the sentences. And the children's songs in the entire corpus are used for training statistical machine learning.

5.2　Experimental Design and Results Analysis

Feasibility of Replacement of the "Theory of Lyrics of Songs" with Statistical machine Learning. In order to verify that the statistical machine learning can replace the "Theory of Lyrics of Songs", the manual classification auxiliary song creation dictionary, in this paper, 15 groups of key words are selected from the "Theory of Lyrics of Songs", to expand the lyrics set and generate the first sentence according to the same algorithm by the "Theory of Lyrics of Songs" and the main body model, respectively. The generated results are evaluated by six reviewers, if the "Theory of Lyrics of Songs" is superior to the statistical machine learning, it will get 1 point; otherwise, 0 point. And the arithmetic mean is taken in the end. The evaluation results are shown in Table 3.

Table 3.　Comparison of lyrics extension methods

Lyrics Set Extension Method	"Theory of Lyrics of Songs"	Statistical machine Learning
Manual Scoring Mean Value	0.4625	0.5375

It can be seen from Table 3 that, it is feasible to replace the "Theory of lyrics in songs" with the statistical machine learning for the expansion of the lyrics words. This not only has broken the constraints in the selection of key words, but from the evalua-

tion results, the main model lyrics expansion single sentence generation result is superior to the lyrics generation result according to the "Theory of lyrics of songs".

Manual Evaluation of Generated Children's Songs. Considering that the current evaluation of the quality of works of art machinery is mainly carried out through the Turing test approach, this paper adopts the expert group manual evaluation method. The evaluation criteria mainly refer to some contemporary scholars' opinions on the evaluation criteria of the children's songs. The reference books include "Twenty-four Songs" by Situ Kong, "Liuyi Songs" by Yang Xiu and "Baishi Taoist Songs" by Jiang Kui. Although these masters have their own evaluation criteria on the quality of the children's songs, in general, the following requirements shall be met: Fluent language, beautiful rhythm, and clear theme, consistent expression of content and with artistic conception. Accordingly, five criteria of manual evaluation are established: Language fluency, rhythm coincidence degree, theme relevance, content expression consistency and artistic conception.

In this paper, the children's songs generation system of He et al. [12] is used as the benchmark system, and the children's song generation system is added to the main model in reference to the benchmark system. In order to get objective evaluation results, the children's songs used by the experimental part is not generated by human interaction, but by using the automatic generation method: Firstly the first sentence is generated by the theme, and then the next sentence is automatically generated, until the entire song is completed. The evaluation data includes 20 children's songs, which are generated according to different theme words, with 10 five character songs and seven character songs each, which are randomly selected from 50 five character songs and 50 seven character songs generated by two different systems, respectively. The single sentence evaluation is to evaluate the first sentence extracted from the 20 songs, □□□the preceding and next sentence evaluation is to randomly select the first and second, second and third, or third and fourth sentence pair from the 20 songs for evaluation. Six evaluators with the knowledge of the field of children's songs evaluation score according to the evaluation criteria. The scoring is conducted according to 1 (very poor) to 5 (excellent) for the three different aspects including single sentence generation, preceding and next sentence generation and the entire song. The experimental results are shown in Table 4, the top data in the single row of the table is the evaluation result of the benchmark system, and the bottom data is the score for the new system.

From the results of the single sentence generation, it can be seen that the system after adding the theme relevance has great improvement on the related performance of the theme. The language fluency is also greatly enhanced compared with the benchmark system due to the consistency of the content of the expression. With the increase in the amount of information between the preceding and next sentence, the evaluation data in each data has declined compared to a single sentence, but from the final manual evaluation results, the system output is acceptable. In the perspective of the evaluation of the entire song, due to the increase in content, the consistency and other aspects has certain decrease. Overall speaking, the new system is superior to the benchmark system.

Table 4. Manual evaluation results

Evaluation Dimension	Single Sentence	Preceding and Next Sentence	The Entire Song
Language Fluency	4.062 4.275	3.575 4.004	3.525 3.912
Prosodic Compliance	— —	3.796 4.096	3.658 3.992
Theme Relevance	3.942 4.250	3.663 3.825	3.408 3.896
Content Consistency	3.983 4.225	3.600 3.788	3.433 3.667
Conception	3.921 3.996	3.563 3.863	3.417 3.883

6   Conclusion

At present, most researches on song generation at home and abroad are based on the template generation method. And in this paper, the original method of combining the statistical machine learning and statistical machine learning translation model is proposed to carry out the automatic generation of the children's songs. The relationship between preceding and next sentences in children's songs is mapped into the bilingual relationship in the statistical machine learning translation. The relevant theories and technical problems under this model are discussed in depth, and the automatic generation of children's songs under this principle is realized. Finally, the development of a strict standard for single sentence generation, the generation of the current sentence according to the preceding context and automatic generation of the whole song is evaluated manually. From the evaluation results, this study has achieved good results, and both the automatic generation of children's songs and natural language has certain reference value. In the future, it requires more consideration of how to design the layout, so that the contents of a song can have better performance in the coherence of the contents and consistency in the logical level. In addition, the style classification of the songs and the generation of the stylized songs is also worthy of study, that is, how to simulate the thinking of the songwriters according to their different writing style with machine to create songs.

7   References

[1] Papineni K , Roukos S , Ward T , Zhu Wei – Jing . BLEU : A method for automatic evaluation of machine translation// Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics(ACL). Philadelphia , USA , 2002 : 311-318

[2] Lutz T . Stochastische texte . Augenblick , 1959 , 4(1):3-9

[3] Manurung H.Chart generator for rhythm patterned text// Proceedings of the 1st International Workshop on Literature in Cognition and Computer.Tokyo,Japan,1999:15-19

[4] Gervás P. WASP:Evaluation of different strategies for the automatic generation of spanish verse// Proceedings of the AISB00Symposium on Creative & Cultural Aspects and Applications of AI &Cognitive Science.Birmingham,UK,2000:93-100

[5] Oliveira H G,Cardoso F A,Pereira F C.Tra-la-Lyrics: An approach to generate text based on rhythm//Proceedings of the 4th International Joint Workshop on Computational Creativity. London, UK,2007:47-55

[6] Kempe V, Levy R, Graci C. Neural networks as fitness evaluators in genetic algorithms:Simulating human creativity// Proceedings of the 23rd Annual Conference of the Cognitive Science Society. Edinburgh,Scotland,2001:1221

[7] Manurung H.An Evolutionary Algorithm Approach to Poetry Generation [Ph. D. dissertation]. University of Edinburgh, Edinburgh,UK,2003

[8] Gervás P.An expert system for the composition of formal Spanish poetry. Journal of Knowledge-Based Systems, 2001,14(3-4):181-188 https://doi.org/10.1016/S0950-7051(01)00095-8

[9] Díaz-Agudo B,Gervás P,González-Calero P A.Poetry generation in colibri//Proceedings of the 6th European Conference on Advances in Case-Based Reasoning. London, UK,2002:73-102

[10] Zhou Chang-Le, You Wei, Ding Xiao-Jun. Genetic algorithm and its implementation of automatic generation of Chinese SONGCI. Journal of Software,2010,21(3):427-437(in Chinese) https://doi.org/10.3724/SP.J.1001.2010.03596

[11] Jiang Long , Zhou Ming . Generating Chinese couplets using a statistical MT approach// Proceedings of the 22nd International Conference on Computational Linguistics. Manchester, England, 2008: 377-384

[12] He Jing , Zhou Ming , Jiang Long . Generating Chinese classical poems with statistical machine translation models// Proceedings of the 26th AAAI Conference on Artificial Intelligence . Toronto, Canada, 2012:1650-1656

[13] Genzel D , Uszkoreit J , Och F. "Poetic" statistical machine translation: Rhyme and meter// Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. Massachusetts, USA , 2010:158-166

[14] Och F J. Minimum error rate training in statistical machine translation//Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics(ACL 2003). Sapporo, Japan, 2003:160-167 https://doi.org/10.3115/1075096.1075117

[15] Koehn P, Och F J, Marcu D. Statistical phrase-based translation// Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology(HLT/NAACL). Edmonton,Canada,2003:48-54 https://doi.org/10.21236/ADA461156

8   Author

Li Pan is with the College of music and dance, Mudanjiang Normal College, Mudanjiang, China (panli_1982@163.com).