ISSN: 0975-766X
CODEN: IJPTFI
Available Online through *Research Article*
www.ijptonline.com

AN INVESTIGATION OF GRAMMARS AND THEIR APPLICATIONS IN RNA STRUCTURE PREDICTION

K.Senthil Kumar, D.Malathi*, J. D. DorathiJayaseeli
Assistant Professor, Department of Computer Science and Engineering, SRM University, Kattankulathur – 603 203.
Professor, Department of Computer Science and Engineering, SRM University, Kattankulathur – 603 203.
Assistant Professor, Department of Computer Science and Engineering, SRM University, Kattankulathur – 603 203.
Email: malathi.d@ktr.srmuniv.ac.in

Received on 20-08-2016 *Accepted on 25-09-2016*

Abstract:

DNA(Deoxyribonucleic acid)andRibonucleic acid (RNA)known as nucleic acids combined with proteins and carbohydrates determine all known form of life that exists in the world.They are the basic units in Genes which do various biological roles like coding, decoding, regulation, and expression of Genes. Some RNA molecules do communication with other cells,controls gene expression, supporting biological reactions. So understanding the functionality of RNA and predicting the structure of RNA is a paramount task in Biotechnology and Genetic Engineering. Grammars play a prominent role in Language theory.Its application includes Computer Vision, Pattern Recognition, Image Processing,Software Engineering etc. In this paper we have analyzedrole of various grammars mainly in RNA structure prediction and its applications in various Genetic Engineering field.

Key words: DNA, RNA, Context Free Grammar,Pseudo knots, Protein.

1. Introduction

DNA and RNA are the important cell molecules for carrying genetic instructions. They are responsible for the growth, development and functioning of all known living organisms and many viruses. In DNA, each nucleotide is composed ofeither cytosine (C), guanine (G), adenine (A), or thymine (T). As per the base pairing rules A with T, and C with G, form adouble-stranded DNA.In RNA, each nucleotide contains adenine (A), cytosine (C), guanine (G), or uracil (U).In RNA, we have single strand folded ontoit. It is known that Adenine and Guanine are purines, cytosine and uracil are pyrimidine.

We have made a thorough analysis in the literature to account for the use of Grammars in various RNA structure predictions. Since grammars are important constituent in computability theory,where Languages are represented by Grammars,people started using grammars in many field like Computer Vision[Yan Dang *et al.,* 2005],[YasuoUemura*et al.,* 1999],[Song-Chun Zhu*et al.,* 2007], Image Processing activities like Segmentation[Hong Chen*et al.,* 2006][Brandon Rothrock*et al.,* 2011],Object Recognition[Long (Leo) Zhu*et al.,* 2012],[ZhuowenTu*et al.,*2005],Object Detection[Feng Han*et al.,* 2009].

Activity Recognition[Yibiao Zhao*et al.,* 2013][Hanumantha T Reddy*et al.,* 2009],Information extraction[M.S.Ryoo*et al.,* 2006][M. Daldoss*et al.,* 2010],Software Testing[Francisco Alvaro*et al.,* 2011][AtsuhiroTakasu, 2008]etc.So an investigation is needed at this point of time to deal grammars role in RNA structure prediction. This detailed survey will form a basis for researchers who will do research in Bio-Informatics where RNA structure can be computed using Grammars.

The literature survey includes the following three steps.

1. Planning
2. Selection
3. Analysis

In the planning phase we have raised following questions.

- What are the problems existing in RNA structure prediction?
- How various algorithms in computer science are used to determine RNA structure prediction?
- How grammars are used in predicting RNA structures?
- What are the other statistical models used for RNA structure prediction?

In the selection phase, we used various data bases, like ACM,Nucleic Acid Research, Journal of Bio-Informatics, Bulletin,IEEE explore etc., to collect relevant papers. Finally we represent a table which comprises of all the reviewed papers.

We organize our paper as follows. Section 2 discusses some preliminary ideas about RNA and its properties. Section 3 deals an over view of grammars, Section 4 deals analysis of papers, section 5 discussion and finally conclusion is given in section 6.

2. Ribo Nucleic Acid(RNA)

Ribo Nucleic Acid is defined as "A nucleic acid present in all living cells and many viruses, consisting of a long usually single stranded chain of alternating phosphate and ribose units, with one of the bases Adenine(A),Guanine(G),Cytosine(C) and uracil (U)".According to [Hu Li*et al.*, 2004],RNA is merely a passive carrier of genetic information between DNA and protein.

2.1 RNA Structure

Ribonucleic acid (RNA) is a biopolymer macromolecule as DNA. It consists of small subunits called nucleotides composed of:

- Purine nucleobases [Adenine−(A), Guanine−(G)]
- Pyrimidine nucleobases [Cytosine−(C), Uracil−(U)]
- D-ribose pentose sugars [$C_5H_{10}O_5$]
- Phosphate groups [PO_4^{3-}]

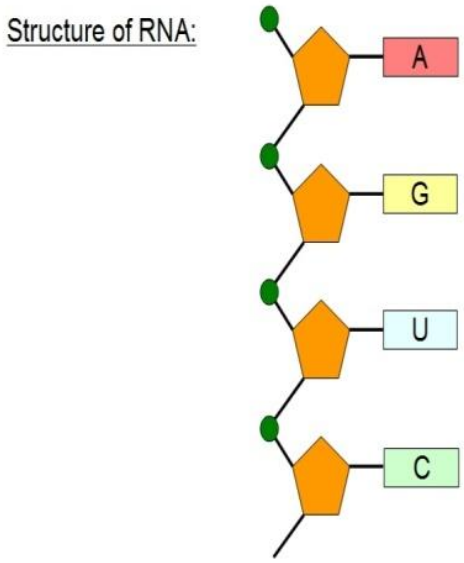

Fig-1:Structure of RNA.

RNA Structure is shown in the Fig-1. The nucleobase is attached on the D−ribose by an N−glycoside bond. The ribose is bonded to the phosphate group through ester bonds. The backbone bonding between RNA nucleotides (i.e. the bonds between the phosphate group and an adjacent ribose sugar) occurs through phosphodiester bonds. A phosphate group is attached to the 3'−carbon position of one ribose and on the 5'−carbon position of the next.

2.2 Various Types of RNA

There are four types of RNA, each encoded by its own type of gene:

1.mRNA - Messenger RNA: It encodes amino acid sequence of a polypeptide.

2.tRNA- Transfer RNA: It brings amino acids to ribosomes during translation

3.rRNA - Ribosomal RNA: With ribosomal proteins, makes up the ribosomes, the organelles that translate the mRNA.

4.snRNA - Small nuclear RNA: With proteins, forms complexes that are used in RNA processing in eukaryotes. (Not found in prokaryotes.)

2.3Definition: RNA Structure Prediction

RNA secondary structure prediction is the process of predicting the position of hydrogen bonds in an RNA molecule based only on its nucleotide sequence. These predictions can be used to better understand the functioning of cells, characteristics of gene expression and the mechanisms involved in protein production. According to [John E. Hopcroft*et al.,* 2001] the folding of a single stranded RNA molecule is determined by its nucleotide sequence. Fig-2 represents the formation of base pairs like A-U, G-C and G-U gives rise to specifics structural motifs like double helical or stem regions and single stranded regions like hairpin-, bulge-, multi branched- and interior loops. The ensemble of these structural elements in a planar presentation is called the secondary structure of RNA.

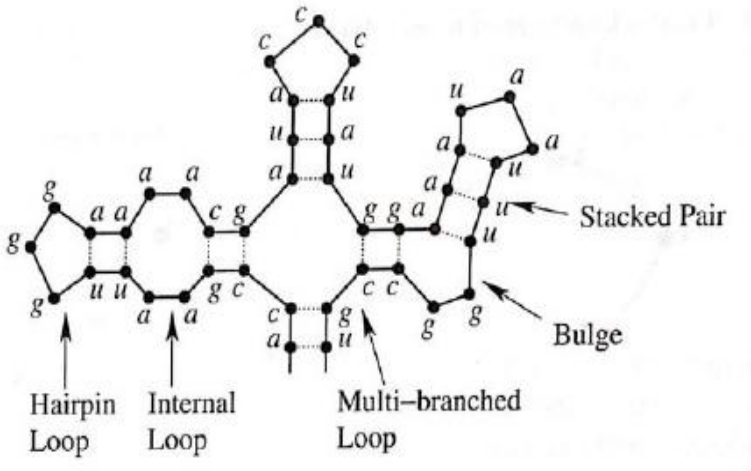

Fig-2: Formation of base pairs like A-U, G-C.

2.4Two important methods for RNA Secondary structure predictions

i) Co variation analysis/Comparative sequence analysis

(This considers the patterns of base pairs during evolution and checks the structural integrity) Molecules with similar functions and different nucleotide sequences will form similar structures. This is based on the assumption that underlying sequence will determine secondary (as well as tertiary) structure. This method helps to find majority of RNA secondary structures

ii) Minimum Free-Energy Method

(By analyzing one sequence we can find other complementary regions which are stable with respect to energy levels)

The method works as follows.

Given RNA linear structure: $R=r_1\ r_2\ .\ .\ .\ r_k$ from *{A, C, G, U}.* To find RNA secondary structure, pairs (r_i,r_j) such that $0<i<j<k+1$ with the condition of secondary structure with possible minimum free energy.One should use "*Tinoco-Uhlenbeck postulate*" which says the energy of each base pair is independent of all of the other pairs and the loop structure. This proves that total free energy is the sum of all of the base pair free energies. Dynamic Programming [M. Zuker, 1989] is the tool to solve the above, but the presence of pseudo knot will cause problem as it breaks recurrence relation.

2.5 Secondary structure without pseudoknots

According to [James WJ Anderson*et al.,* 2012] Secondary structure without pseudoknots is defined as follows.

Definition: Secondary structure without pseudoknots

Let $A = a_1a_2$: : : an be an RNA sequence. That is, A is a string over an alphabet $\sum$= *{a; u; g; c}.* A pair of residues (letters) (*x; y*) is called a (complementary) basepair if *{*x; y*} = {*a; u*}* or *{*x; y*} = {*g; c*}.* Although the wobble pair *{g; u}* is nottreated as a base pair, similar results hold for such a case. A set of pairs of indices$M = \{(i; j)|\ 1\leq i<j\leq n,\ 1<i<j<n;$ (a_i; a_j) is a base pair*}* is called an RNA secondary structure without pseudo knots if no distinct pairs (a_i; a_j);(a_h; a_k) in *M* satisfy $i\leq h\leq j\leq k$.

2.5 Secondary Structure with Pseudoknots

2.5.1 Definition: simple pseudoknots

Consider a consecutive subsequence $a_{i0}a_{i0+1}$: : : a_{k0} of an RNA sequence *A* where $i_0$and $k_0$are arbitrarily chosen positions. We call a set of base pairs $M_{i0;\ k0}$ a simple pseudo knot if there exist positionsj_0; $j_{0'}(i_0<j_{0'}<j_0<k_0)$ for which the following conditions are satisfied:

1.Each $i(i_0<i<k_0)$ appears at most once in $M_{i0}; k_0$.

2. Each $(i; j)\in M_{i0}; k_0$satisfies either$i_0\leq i\leq j_0"\leq j<j_0$ or $j_0"\leq i\leq j_0\leq k_0$.

3.If pairs $(i; j)$ and $(I'; j')$ in $M_{i0}; k_0$satisfy either $i<I'<j_0'$ or $j_0"\leq i<I"$, then $j>j_0$holds.

2.5.2 Definition:RNA secondary structure with simple pseudoknots

A set of base pairs *M* is called an RNA secondary structure with simple Pseudoknotsif the following conditions are satisfied:

1.$M=M'UM_{i1,k1}UM_{i2,k2}U\ldots\ldots M_{it,kt}$,twhere *t* is a non-negative integer and $1\leq i_1<k_1<i_2<k_2<\ldots.<i_t<k_t\leq n$.

2.Each $M_{ih}; k_h$is a simple pseudoknot for a consecutive subsequence $a_{ih}a_{ih+1} : : : a_{kh}$.

2.6 Some more methods to predict RNA structures

According to [Jeff Offutt*et al.*, 2006] following methods are used to determine the structure of RNA.

1.Sequence Alignment (The ability of bases to align with other bases)

2. Base Pair Maximization (using Dynamic Programming)

3. Binary Tree Representation (Representing RNA structure as a Binary tree form)

4. Covariance Model

3. An Overview of Grammars

Grammars are basic building blocks to represent a Language. Their recursive nature helps us to generate various strings of a language. As in [Matthew G. Seetin et al, 2012], a Grammar is defined as $G=(V,T,P,S)$

Where *V*-denotes a finite set of variables, *T*-denotes a finite set of terminals, *P*-denotes a finite set of production rules and *S*-denotes a special start symbol.

3.1 Chomsky Hierarchy of Grammars

According to Chomsky[Tatsuya Akutsu, 2000] grammars can be classified as follows.

1. *Type-0* grammars (unrestricted grammars) include all formal grammars
2. *Type-1* grammars (context-sensitive grammars). These grammars have rules of theform $\alpha A\beta$--> $\alpha\gamma\beta$-with *A*, a non-terminal and,and α,β,γstrings of terminals and/or non-terminals.
3. *Type-2* grammars (context-free grammars). Here rules are of the form $A\rightarrow\alpha$ where*A* is a variable and $\alpha\in(VUT)^*$

4. *Type-3* grammars (regular grammars).Here production rules are of the form $A \rightarrow Ba$ or $A \rightarrow aB$,or $A \rightarrow a$.

3.2 Stochastic Context Free Grammar

Definition:A Stochastic Context Free Grammar is defined as $G=\{V,T,P,S,R\}$,Where$V$ is the set of non-terminal symbols, T is the set of terminal symbols, P is the set of production rules, S is the start symbol and R is the set of probabilities on production rules.

Analysis of Survey papers: We have analyzed the various papers that were selected in selection phase;We mainly considered the articles which use stochastic grammar and various statistical model for RNA structure prediction. The Analysis of the papers are shown in the Table-1.

Table-1: Analysis of papers.

Sl. No	Authors	Objective	Algorithm /Techniques	Grammar Used	Data set /Tool Kit	Result
1	James WJ Anderson et al 2012	RNA secondary structure prediction	Grammars in Chomsky Normal Form, CYK Algorithm, Inside-Outside algorithm and Evolutionary algorithm	Stochastic Context Free Grammar	RNASTRAND data set, The tmRDB and SRPDB resource, The nucleic acid database, The Ribonuclease P Database.	Sensitivity and specificity of evolved SCFGs and other prediction methods are presented in a table
2.	Elena Rivasetal2001	Noncoding RNA gene detection	SCFG, pair-HMMs	Stochastic Context Free (nine different set of SCFG) Grammar	QRNA , CRITICA, EXOFISH	Generatedpair wise alignments of about 40% sequence identity
3	Bjarne Knudsen Etal 2003	Pfold: RNA secondary structure prediction	The KH-99 algorithm	Stochastic Context Free Grammar	An alignment of up to 40 sequences and 500 positions are used	Evaluating the prediction accuracy as a function of the number of sequences used in the analysis are shown in a table
4.	RuanJetal 2004	RNA secondary structure prediction	Iterated loop matching approach	Grammar is not specified	Algorithm is tested on a Number of RNA families.	The algorithm finds 90% of base-pairs for short sequences and 80% overall
5	Eddy SR etal 2000	RNA secondary structure prediction	Probabilistic Models of RNA Secondary Structure Including Pseudoknot	RNA Pseudoknot Grammar	Not specified	One-One correspondence between RNA structure prediction using Dynamic programming and Proposed algorithm is shown
6	Sean R. Eddy etal	RNA structure prediction	Dynamic Programming	Grammar Not specified	A Silicon Graphics	A 100 nucleotide RNA takes about 4 hours and

Sl. No	Authors	Objective	Algorithm /Techniques	Grammar Used	Data set /Tool Kit	Result
					Origin200	22.5 MB to fold on an SGI R10K Origin200
7.	Jan Pieter Abrahams1 Etal 1998[21]	RNA secondary structure Prediction	Present a computer program	Grammar not specified	Amdahl V7B mainframecomp uter with an APL SV 4.0 interpreter system.	Comparison of the secondary structures of tRNA molecules as predicted by various programs are shown in a table
8	Tianhua Wu etal 2009[22]	RNA Secondary Structure Prediction	Bayesian network,Zuker algorithm and SCFG.	Stochastic Context Free Grammar used	tRNA dataset, EMBL databank, EUBACT.	The correlation coefficient, Sensitivity, Specificity values of proposed method with existing methods are shown in a table
9	Tatsuya Akutsu 2000[23]	RNA Secondary Structure Prediction	Dynamic programming	Grammar not specified	Not specified	Dynamic algorithm results shown in a table.
10	YasubumiSak akibara 2005[24]	RNA Secondary Structure Prediction	Pair Hidden Markov Models and SCFG	SCFG specified	EUBACT, CYACHL,ARC HAE	Prediction accuracy of tRNA is 100%, 99.87%,99.86% respectively with respect to the given data set

5. Discussion

According to[Robin D Dowell *et al.*, 2004], there are several fundamental tasks in biological sequence analyses:

1. Calculating multiple alignments
2. Representing motifs
3. Classifying proteins and predicting their structures and functions
4. Providing unifying frameworks for various gap models for pairwise alignments
5. Finding gene encoding regions on genome sequences
6. Analysing promoter regions and their transcription regulations, and
7. Identifying signal sequences for protein localizations and interactions.

We have gone through the literature to see the role of stochastic grammars and how they used to fulfil one or other tasks in predicting RNA structure.

In this survey we have reviewed papers which use grammars for the various tasks. [James WJ Andersonetal. 2012**]** used Chomsky Normal Form of grammar for RNA structure prediction. Two automatic search techniques for effective grammars – exhaustive search for very compact grammars and an evolutionary algorithm to find larger grammars are

proposed by them.Robin D Dowell etal. [Robin D Dowell *et al.,* 2004] describes a comparative sequence analysis algorithm for detecting novel structural RNA genes. The key idea is to test the pattern of substitutions observed in a pair wise alignment of two homologous sequences. A conserved coding region tends to show a pattern of synonymous substitutions, whereas a conserved structural RNA tends to show a pattern of compensatory mutations consistent with some base-paired secondary structure.

According to Bjarne Knudsen etal. [Bjarne Knudsen*et al.,* 2003],a Stochastic Context-Free Grammar (SCFG) is used to present a prior probability distribution of RNA structures. Given an alignment and a phylogenetic tree relating the sequences, posterior probabilities of the structures can be calculated using the inside–outside algorithm. The posterior probability is based on individual probabilities for alignment columns or pairs of columns in the case of a base-pair.

Ruan J etal. 2004[Ruan J*et al.,* 2004] describe an algorithm, known as iterated loop matching, for reliable and efficient prediction of RNA secondary structures including pseudo knots. The method can utilize either thermodynamic or comparative information or both, thus is able to predict pseudo knots for both aligned and individual sequences. Using 8-12 homologous sequences, the algorithm correctly identifies more than 90% of base-pairs for short sequences and 80% overall.

Rivas E, etal. [Rivas E*et al.,* 2000]present grammatical representations of RNA structure with Pseudo knots. The grammars are able to generate various Pseudoknot structures: Elena Rivas etal. [Elena Rivas *et al.,* 1999] use a dynamic programming algorithm for predicting optimal RNA secondary structure, including pseudoknots. The algorithm has a worst case complexity of $O(N^6)$ in time and $O(N^4)$ in storage.[Jan Pieter *et al.,* 1990]shows a program which is able to predict pseudo knotted structures together with a secondary structure.

[TianhuaWu *et al.,* 2009] provides an effective method of combining free energy information of Zuker algorithm with statistical information from SCFG probability model to predict RNA secondary structure. [Tatsuya Akutsu2000] presented a simple dynamic programming algorithm for RNA secondary structure prediction with pseudoknots.[YasubumiSakakibara2005]uses Stochastic grammar for the biological sequence analysis.

5.2. Various Models to predict RNA Secondary Structure

Table-2 shows various models which analyse RNA Structure.[Krogh A et al. 1994] represent a statistical model viz Hidden Markov Model to database searching and multiple sequence alignment of protein families and protein domains.

[Eddy SR *et al.,* 1994] proposed dynamic programming algorithm for predicting optimal RNA secondary structure, including pseudo knots.

[Felsenstein. J et al. 1981] prescribed maximum likelihood techniques to the estimation of evolutionary trees which are constructed nucleic acid sequence data. [E Rivas *et al.,* 2001]describe a comparative sequence analysis algorithm for detecting novel structural RNA genes. They proposed a method which used to i test the pattern of substitutions observed in a pair wise alignment of two homologous sequences. [YasubumiSakakibara*et al.,* 1004] present an algorithm, iterated loop matching, for reliably and efficiently predicting RNA secondary structures including pseudo knots.The method can utilize either thermodynamic or comparative information or both, thus is able to predict pseudo knots for both aligned and individual sequences.

[Tatsuya Akutsu, 2000] shows simple dynamic programming algorithms for RNA secondary structure prediction with pseudo knots.

[YasuoUemura, 1999] proposed a subclass of tree adjoining grammars (TAGs) that is suitable for the application to modelling and predicting RNA secondary structures

Table-2 Various Models to predict RNA Secondary Structure

Sl. No	Authors	Objective	Algorithm /Techniques	Grammar Used	Data set /Tool Kit	Result
1	James WJ Anderson et al 2012	RNA secondary structure prediction	Grammars in Chomsky Normal Form, CYK Algorithm, Inside-Outside algorithm and Evolutionary algorithm	Stochastic Context Free Grammar	RNASTRAND data set, The tmRDB and SRPDB resource, The nucleic acid database, The Ribonuclease P Database	Sensitivity and specificity of evolved SCFGs and other prediction methods are presented in a table
2.	Elena Rivasetal 2001	Noncoding RNA gene detection	SCFG, pair-HMMs	Stochastic Context Free (nine different set of SCFG) Grammar	QRNA , CRITICA, EXOFISH	Generated pair wise alignments of about 40% sequence identity
3	Bjarne Knudsen	Pfold: RNA	The KH-99 algorithm	Stochastic Context Free	An alignment of up to 40 sequences and 500	Evaluating the prediction accuracy as a

	Etal 2003	secondary structure prediction		Grammar	positions are used	function of the number of sequences used in the analysis are shown in a table
4.	RuanJetal 2004	RNA secondary structure prediction	Iterated loop matching approach	Grammar is not specified	Algorithm is tested on a Number of RNA families	The algorithm finds 90% of base-pairs for short sequences and 80% overall
5	Eddy SR etal 2000	RNA secondary structure prediction	Probabilistic Models of RNA Secondary Structure Including Pseudoknot	RNA Pseudoknot Grammar	Not specified	One-One correspondence between RNA structure prediction using Dynamic programming and Proposed algorithm is shown
6	Sean R. Eddy etal	RNA structure prediction	Dynamic Programming	Grammar Not specified	A Silicon Graphics Origin200	A 100 nucleotide RNA takes about 4 hours and 22.5 MB to fold on an SGI R10K Origin200
7.	Jan Pieter Abrahams1 etal 1998[21]	RNA secondary structure Prediction	Present a computer program	Grammar not specified	Amdahl V7B mainframe Computer with an APL SV 4.0 interpreter system.	Comparison of the secondary structures of tRNA molecules as predicted by various programs are shown in a table
8	Tianhua Wu etal 2009[22]	RNA Secondary Structure Prediction	Bayesian network,Zuker algorithm and SCFG.	Stochastic Context Free Grammar used	tRNA dataset, EMBL databank, EUBACT.	The correlation coefficient, Sensitivity, Specificity values of proposed method with existing methods are shown in a table
9	Tatsuya Akutsu 2000[23]	RNA Secondary Structure Prediction	Dynamic programming	Grammar not specified	Not specified	Dynamic algorithm results shown in a table.
10	YasubumiS	RNA	Pair Hidden Markov	SCFG	EUBACT, CYACHL,	Prediction accuracy of

	akakibara 2005[24]	Secondary Structure Prediction	Models and SCFG	specified	ARCHAE	tRNA is 100%, 99.87%,99.86% respectively with respect to the given data set

6. Conclusion

Predicting RNA structure with Pseudo knot presence is a challenging one. Many people tried varies methods like Dynamic Programming method, Zuker algorithm, Hopfield network, Local Search methods, Tree adjoining Grammar, Profile HMMs Pair HMMs, etc. But these methods are not guaranteed to give optimal structures. When simple Pseudo knot occurs these methods may find solution, but when complex Pseudo knot occurs predicting RNA structures are always very difficult. So, there are lot of research scope one can find in this field. Combining grammatical framework with efficient algorithms in computer Science may lead potential success in RNA Structure Prediction. From the analysis one can find that more work needs to be done in Biological Sequence Analysis with respect to Grammar point of view. Effective use of various Grammars can lead to a dominant research area in RNA Structure Prediction.

References

1. Yan Dang, Yulei Zhang, Dong-mo Zhang (2005),Statistical Parser For RNA Secondary Structure Prediction, *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, Vol 18,* No.21,pp3399-3403.
2. YasuoUemura, Aki Hasegawa, Satoshi Kobayashi, Takashi Yokomori(1999),Tree adjoining grammars for RNA structure prediction, *Theoretical Computer Science, Vol 210,*pp277-303.
3. Song-Chun Zhu, David Mumford (2007),A stochastic Grammar of Images, *Foundations and Trends in Computer Graphics and Vision, Vol 2,* pp259-362.
4. Hong Chen, ZiJianXu, ZiQiang Liu, Song Chun Zhu (2006), Composite Templates for Cloth Modeling and Sketching, *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp943-950.
5. Brandon Rothrock, Song-Chun Zhu(2011), Human parsing using stochastic AND-OR grammars and rich appearances, *ICCV Workshops*, pp640-647.

6. Long (Leo) Zhu, Yuanhao Chen, Yuan Lin, Chenxi Lin, Alan Yuille (2012), Recursive Segmentation and Recognition Templates for Image Parsing, *IEEE transactions on pattern analysis and machine intelligence, Vol 34*, pp359-371.
7. ZhuowenTu, Xiangrong Chen, Alan L. Yuille, and Song-Chun Zhu (2005), Image Parsing: Unifying Segmentation, Detection, and Recognition, *Int'l J. of Computer Vision, Marr Prize Issue*, pp1-45.
8. Feng Han, Song-Chun Zhu (2009), Bottom-up/Top-Down Image Parsing by Attribute Graph Grammar, *ICCV05, Vol 2*, pp1778 - 1785.
9. Yibiao Zhao, Song-Chun Zhu (2013), Scene Parsing by Integrating Function, Geometry and Appearance Models, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp3119-3126.
10. Hanumantha T Reddy, K. Karibasappa, A. Damodaram (2009), Detection on Methods and Models in Computer Science, *ICM2CS'09*, pp1-7.
11. M.S.Ryoo, J. K. Aggarwal (2006), Semantic Understanding of Continued and Recursive Human Activities, *18th International Conference on Pattern Recognition (ICPR'06), Vol1*, pp378-379.
12. M. Daldoss, N. Piotto, N. Conci, F.G.B. DeNatale (2010), Learning and matching human activities using regular expressions, *IEEE International Conference on Image Processing*, pp4681-4684.
13. FranciscoAlvaro, Joan-AndreuS´nchez, Jose-Miguel Benedi (2011), Recognition of Printed Mathematical Expressions Using Two-Dimensional Stochastic Context-Free Grammars, *International Conference on Document Analysis and Recognition*, pp1225-1229.
14. AtsuhiroTakasu (2008), Information Extraction by Two Dimensional Parser, *20th IEEE International Conference on Tools with Artificial Intelligence*, pp333-340.
15. Hu Li, Maozhong Jin, Chao Liu, ZhongyiGao (2004), Test Criteria for Context Free Grammars, *Proceedings of the 28th Annual International Computer Software and Applications Conference (COMPSAC'04)*.
16. Jeff Offutt, Paul Ammann, Lisa (Ling) Liu (2006),Mutation Testing implements Grammar-Based Testing, *Proceeding Mutation '06 Proceedings of the Second Workshop On Mutation Analysis, IEEE Computer society Washington DC.*

17. Matthew G. Seetin, David H. Mathews (2012), RNA Structure Prediction: An Overview of Methods, Bacterial Regulatory RNA, *series Methods in Molecular Biology, Vol 905*, pp99-122.
18. Jan Pieter Abrahams, Mirjam van den Berg, Eke van Batenburg and CornelisPleij(1990),Prediction of RNA secondary structure including pseudo knotting by computer simulation, *Nucleic Acids Research, Vol 18*, No. 10, pp3035-3044.
19. Tatsuya Akutsu (2000), Dynamic programming algorithms for RNA secondarystructure prediction with pseudoknots, *Discrete Applied Mathematics,Vol 104*, pp45-62.
20. Eddy S.R (2004), How Do RNA Folding Algorithms Work?*Nature Biotechnology, Vol 22*, pp1457-1458.
21. John E. Hopcroft, Rajeev Motwani, Jeffrey D. Ullman (2001), *Introduction to Automata Theory, Languages, and Computation, Addison-Wesley*.
22. Chomsky, Noam(1959),Three models for the description of language, *IRE Transactions on Information Theory Vol 2*, pp113–124.
23. James WJ Anderson, Paula Tataru, Joe Staines, Jotun Hein and Rune Lyngs (2012), Evolving stochastic context-free grammars for RNA secondary structure prediction**,** *BMC Bioinformatics*, pp13-78.
24. Robin D Dowell and Sean R Eddy (2004), Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction, *BMC Bioinformatics Vol 5*, No. 71.
25. Bjarne Knudsen and JotunHei (2003), Pfold: RNA secondary structure prediction using stochastic context-free grammars, *Nucleic Acids Res. Vol 31*, No.13, pp3423–3428.
26. Ruan J, Stormo GD, Zhang W. (2004), An iterated loop matching approach to the prediction of RNA secondary structures with pseudo knots**,** *Bioinformatics,Vol 20*, No. 1, pp58-66.
27. Rivas E, Eddy SR. (2000), The language of RNA: a formal grammar that includes pseudoknots, *Bioinformatics,Vol 16*, No. 4, pp334-340.
28. Elena Rivas and Sean R. Eddy (1999), A dynamic programming algorithm for RNA structure prediction including pseudoknots, *J. Mol. Biol. Vol 285*, pp2053-2068.

29. Tianhua Wu, Zhidong Deng, DandanSong (2009), An approach for RNA secondary structure prediction based on Bayesian network,*IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB '09*, pp 24-30.
30. Tatsuya Akutsu (2000), Dynamic programming algorithms for RNA secondary structure prediction with pseudo knots, *Discrete Applied Mathematics, Vol 104*, pp45-62.
31. YasubumiSakakibara (2005), Grammatical Inference in Bioinformatics, *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 27*, No.7, pp1051-1062.
32. YasubumiSakakibara, Michael Brown, Richard Hughey, I.SairaMian, KimmenSjolander (1994),Stochastic context-free grammars for tRNA modelling,*Nucleic Acids Research, Vol 22*, No.23,pp5112-5120.
33. M. Zuker (1989), The Use of dynamic programming in RNA secondary structure prediction, *Mathematical Methods for DNS Sequences. Boca Raton, FL: CRC Press*.
34. Anders Krogh'f, Michael Brown' I., SairaMian' KiminenSjolander' and David Hausder'S, (1994), Hidden Markov Models in Computational Biology Applications to Protein Modeling, *Journal of Molecular Biology*, pp.1501-1531.
35. Eddy SR, Durbin R (1994), RNA Sequence Analysis Using Covariance Models**,** *Nucl Acids Res, Vol22***,** pp2079–2088.
36. Felsenstein J. (1981), Evolutionary trees from DNA sequences: a maximum likelihood approach, *J. Mol. Evol.,Vol 17*, pp368–376.
37. Rivas .E, Eddy SR (2001), Noncoding RNA gene detection using comparative sequence analysis, *BMC Bioinformatics, Vol 2*, No. 8.