Probability and Recursion

Kousha Etessami[1] and Mihalis Yannakakis[2]

[1] LFCS, School of Informatics, University of Edinburgh
[2] Department of Computer Science, Columbia University

In this talk we will discuss recent work on the modeling and algorithmic analysis of systems involving recursion and probability. There has been intense activity recently in the study of such systems [2,3,10,11,13,14,15,16,17]. The primary motivation comes from the analysis of probabilistic programs with procedures. Probability can arise either due to randomizing steps in the program, or it may reflect statistical assumptions on the behaviour of the program, under which we want to investigate its properties.

Discrete-time, finite state Markov chains have been used over the years in a broad range of applications to model the evolution of a variety of probabilistic systems. Markov Decision Processes are a useful model for control optimization problems in a sequential stochastic environment that combines probabilistic and nonprobabilistic aspects of system behavior [28,19,18]. These models have also been used extensively in particular to model probabilistic programs without procedures and to analyze their properties [7,8,24,27,33]. In the presence of recursive procedures, a natural model for probabilistic programs is *Recursive Markov Chains* (RMCs): Informally, a RMC consists of a collection of finite state component Markov chains that can call each other in a potentially recursive manner [13]. An equivalent model is *probabilistic Pushdown Automata* (pPDA) [10]. These models are essentially a succinct, finite representation of an infinite state Markov chain, which captures the global evolution of the system.

More generally, if some steps of the program/system are probabilistic while other steps are not, but rather are controllable by the designer or the environment, then such a system can be naturally modeled by a *Recursive Markov Decision Process* (RMDP) or a *Recursive Simple Stochastic Game* (RSSG)[15]. In a RMDP all the nonprobabilistic actions are controlled by the same agent (the controller or the environment), while in a RSSG, different nonprobabilistic actions are controlled by two opposing agents (eg. some by the designer and some by the environment).

Some types of recursive probabilistic models arise naturally in other contexts and have been studied earlier, some even before the advent of computer science. *Branching processes* are an important class of such processes [21], introduced first by Galton and Watson in the 19th century to study population dynamics, and generalized later on in the mid 20th century to the case of *multitype branching processes* by Kolmogorov and developed further by Sevastyanov [23,31]. They have been applied in a wide variety of contexts such as population genetics [22], models in molecular biology for RNA [30], and nuclear chain reactions [12]. Another related model is that of *stochastic context-free grammars* which have been

X. Deng and D. Du (Eds.): ISAAC 2005, LNCS 3827, pp. 2–4, 2005.

studied extensively since the 1970's especially in the Natural Language Processing community (see eg. [25]). In a certain formal sense, multitype branching processes and stochastic context-free grammars correspond to a subclass of recursive Markov chains (the class of "single-exit RMCs", where each component Markov chain has a single exit state where it can terminate and return control to the component that called it).

Recursive Markov chains, and their extension to Recursive Markov Decision Processes and Simple Stochastic Games, have a rich theory. Their analysis involves combinatorial, algebraic, and numerical aspects, with connections to a variety of areas, such as the existential theory of the reals [5,29,4], multidimensional Newton's method, matrix theory, and many others. There are connections also with several well-known open problems, such as the square root sum problem [20,32] (a 30-year old intriguing, simple problem that arises often in the numerical complexity of geometric computations, and which is known to be in PSPACE, but it is not known even whether it is in NP), and the value of simple stochastic games [6] and related games, which are in NP∩coNP, but it is not known whether they are in P.

In this talk we will survey some of this theory, the algorithmic results so far, and remaining challenges.

Acknowledgement. Work partially supported by NSF Grant CCF-04-30946.

References

1. R. Alur, M. Benedikt, K. Etessami, P. Godefroid, T. W. Reps, and M. Yannakakis. Analysis of recursive state machines. In *ACM Trans. Progr. Lang. Sys.*, 27:786-818, 2005.
2. T. Brázdil, A. Kučera, and J. Esparza. Analysis and prediction of the long-run behavior of probabilistic sequential programs with recursion. In *Proc. of FOCS'05*, 2005.
3. T. Brázdil, A. Kučera, and O. Stražovský. Decidability of temporal properties of probabilistic pushdown automata. In *Proc. of STACS'05*, 2005.
4. S. Basu, R. Pollack, and M. F. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *J. ACM*, 43(6):1002–1045, 1996.
5. J. Canny. Some algebraic and geometric computations in PSPACE. In *Prof. of 20th ACM STOC*, pages 460–467, 1988.
6. A. Condon. The complexity of stochastic games. *Inf. & Comp.*, 96(2):203–224, 1992.
7. C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the ACM*, 42(4):857–907, 1995.
8. C. Courcoubetis and M. Yannakakis. Markov decision processes and regular events. *IEEE Trans. on Automatic Control*, 43(10):1399–1418, 1998.
9. L. de Alfaro, M. Kwiatkowska, G. Norman, D. Parker, and R. Segala. Symbolic model checking of probabilistic processes using MTBDDs and the kronecker representation. In *Proc. of 6th TACAS*, pages 395–410, 2000.
10. J. Esparza, A. Kučera, and R. Mayr. Model checking probabilistic pushdown automata. In *Proc. of 19th IEEE LICS'04*, 2004.

11. J. Esparza, A. Kučera, and R. Mayr. Quantitative analysis of probabilistic pushdown automata: expectations and variances. *Proc. of 20th IEEE LICS*, 2005.
12. C. J. Everett and S. Ulam. Multiplicative systems, part i., ii, and iii. Technical Report 683,690,707, Los Alamos Scientific Laboratory, 1948.
13. K. Etessami and M. Yannakakis. Recursive Markov chains, stochastic grammars, and monotone systems of non-linear equations. In *Proc. of 22nd STACS'05.* Springer, 2005. (Tech. Report, U. Edinburgh, June 2004).
14. K. Etessami and M. Yannakakis. Algorithmic verification of recursive probabilistic state machines. In *Proc. 11th TACAS*, vol. 3440 of LNCS, 2005.
15. K. Etessami and M. Yannakakis. Recursive Markov Decision Processes and Recursive Stochastic Games. In *Proc. ICALP*, pp. 891-903, Springer, 2005.
16. K. Etessami and M. Yannakakis. Checking LTL Properties of Recursive Markov Chains. In *Proc. 2nd Intl. Conf. on Quantitative Evaluation of Systems*, IEEE, 2005.
17. K. Etessami and M. Yannakakis. Efficient Analysis of Classes of Recursive Markov Decision Processes and Stochastic Games, submitted.
18. E. Feinberg and A. Shwartz, editors. *Handbook of Markov Decision Processes.* Kluwer, 2002.
19. J. Filar and K. Vrieze. *Competitive Markov Decision Processes.* Springer, 1997.
20. M. R. Garey, R. L. Graham, and D. S. Johnson. Some NP-complete geometric problems. In *8th ACM Symp. on Theory of Computing*, pages 10–22, 1976.
21. T. E. Harris. *The Theory of Branching Processes.* Springer-Verlag, 1963.
22. P. Jagers. *Branching Processes with Biological Applications.* Wiley, 1975.
23. A. N. Kolmogorov and B. A. Sevastyanov. The calculation of final probabilities for branching random processes. *Dokl. Akad. Nauk SSSR*, 56:783–786, 1947. (Russian).
24. M. Kwiatkowska. Model checking for probability and time: from theory to practice. In *18th IEEE LICS*, pages 351–360, 2003.
25. C. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing.* MIT Press, 1999.
26. A. Paz. *Introduction to Probabilistic Automata.* Academic Press, 1971.
27. A. Pnueli and L. D. Zuck. Probabilistic verification. *Inf. and Comp.*, 103(1):1–29, 1993.
28. M. L. Puterman. *Markov Decision Processes.* Wiley, 1994.
29. J. Renegar. On the computational complexity and geometry of the first-order theory of the reals, parts I-III. *J. Symb. Comp.*, 13(3):255–352, 1992.
30. Y. Sakakibara, M. Brown, R Hughey, I.S. Mian, K. Sjolander, R. Underwood, and D. Haussler. Stochastic context-free grammars for tRNA modeling. *Nucleic Acids Research*, 22(23):5112–5120, 1994.
31. B. A. Sevastyanov. The theory of branching processes. *Uspehi Mathemat. Nauk*, 6:47–99, 1951. (Russian).
32. P. Tiwari. A problem that is easier to solve on the unit-cost algebraic ram. *Journal of Complexity*, pages 393–397, 1992.
33. M. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Proc. of 26th IEEE FOCS*, pages 327–338, 1985.