Diss. ETH No. 10923

q–Distributions and Random Graphs

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Technical Sciences

presented by

DAVIDE CRIPPA
Dipl. Math. ETH

born October 29, 1964
citizen of Lugano, Switzerland

accepted on the recommendation of

Prof. Dr. K. Simon, examiner
Prof. Dr. A. D. Barbour, co–examiner

1994

Abstract

The theory of graphs is one of those disciplines which has gained more and more importance in the past decades, in particular because of the flexibility and wide range of applications that a graph represents as mathematical structure. This has as consequence the search for more efficient algorithms, especially those that can be applied to the field of real time systems.

"Time utilization" is the complexity measure which is nowadays mostly used to define "efficiency". Within this measure the traditional type of complexity treatment is the "worst–case analysis", although extensive tests of algorithms increasingly show that there is often a considerable discrepancy between the empirical and the theoretical results. This gap could be partly closed by the so–called "average–case analysis". This kind of analysis, though, requires a probability model, which in the theory of graphs is realized by "random graphs".

In this work we have analyzed some graph parameters such as the size of the transitive closure, the number of sources or the width of a greedy chain decomposition of a random graph. We have shown that most of these parameters can be described by the same type of Markov process, which we could analyze under general conditions. In particular we have found closed forms for the distributions and the moments of the corresponding random variables. These results have allowed us to ameliorate the average running time analysis for the computation of the transitive closure and reduction in acyclic digraphs.

The results we have found go well beyond the analysis of algorithms. In particular we point out new relationships between graph theory, probability theory, analytical number theory and the theory of q–hypergeometrical series. The dominating role played by the q–hypergeometrical functions can be directly read by the continuous usage throughout this work of the q–notation, in our eyes a perfect tool to deal with these kinds of problems.

Kurzfassung

Die Graphentheorie ist eine jener Wissenschaften die in den letzten Jahrzehnten vermehrt an Bedeutung gewonnen hat. Einige der Gründe dafür sind gewiss die grosse Flexibilität und Anwendbarkeit, die ein Graph als mathematische Struktur darstellt. Und gerade bei Anwendungen werden immer effizientere Algorithmen verlangt, ganz besonders im Bereich der Echtzeitsystemen.

Der Begriff "Effizienz" wird heute meist mit "Zeitverbrauch" identifiziert. Theoretische Arbeiten zum Zeitverbrauch sind traditionell von der "Analyse im schlechtesten Fall" geprägt, obwohl die erhaltenen Resultate immer öfter von den empirischen Beobachtungen abweichen. Als Alternative dazu entwickelte sich die sogenannte "mittlere Laufzeitanalyse", die aber grundsätzlich ein Wahrscheinlichkeitsmodell voraussetzt. In der Graphentheorie wird dieses durch die "zufälligen Graphen" realisiert.

In dieser Arbeit haben wir einige Graphenparameter analysiert wie die Grösse der transitiven Hülle, die Anzahl Quellen oder die Breite einer Greedy–Wegzerlegung eines zufälligen Graphen. Es zeigte sich, dass sich die meisten dieser Parameter durch einen speziellen Markov–Prozess beschreiben lassen. Diesen Prozess haben wir unter allgemeinen Bedingungen analysiert. Insbesondere haben wir geschlossene Formen für die Verteilungen und Momente der entsprechenden Zufallsvariablen hergeleitet. Diese Resultate erlaubten uns dann, die mittlere Laufzeitanalyse für die Berechnung der transitiven Hülle bzw. der transitiven Reduktion in azyklischen Graphen zu verbessern.

Die Aussagen der vorliegenden Arbeit gehen weit über die Analyse von Algorithmen hinaus. Insbesondere wurden neue Beziehungen zwischen Graphentheorie, Wahrscheinlichkeitstheorie, analytischen Zahlentheorie und der Theorie der q–hypergeometrischen Reihen aufgedeckt. Die dominierende Rolle der q–hypergeometrischen Funktionen lässt sich an dem durchgehenden Gebrauch der q–Notation ablesen, die sich als überaus mächtiges Werkzeug zur Analyse unsere Problemen erwiesen hat.

Riassunto

La teoria dei grafi ha assunto in questi ultimi anni un'importanza sempre maggiore. Uno dei motivi è la flessibiltà offerta da questa struttura matematica e la possibilità di utilizzarla in diversi campi di applicazione. A questo proposito è necessario notare che sono proprio le applicazioni a richiedere algoritmi sempre più efficienti.

Il concetto di efficienza si identifica abitualmente con quello di "tempo impiegato". In generale, nonostante i risultati così ottenuti si discostino spesso da quelli osservati empiricamente, gli studi teorici riguardanti l'efficienza sono basati sull'analisi del caso peggiore (worst–case analysis). Un altro approccio al problema consiste nell'analisi del tempo medio (average–case analysis), analisi che richiede l'uso di un modello probabilistico. Nella teoria dei grafi questo modello è costituito di regola dai "grafi casuali".

In questo lavoro abbiamo analizzato alcuni parametri relativi ai grafi casuali, quali la dimensione della chiusura transitiva, il numero di sorgenti e l'ampiezza di una suddivisione greedy del grafo in percorsi disgiunti. Si può dimostrare che molti di questi parametri si possono descrivere tramite un particolare processo di Markov, che abbiamo analizzato in maniera generale. In particolare abbiamo determinato forme chiuse per le distribuzioni e per i momenti delle corrispondenti variabili casuali. Questi risultati ci hanno permesso di migliorare la stima dei tempi medi per il calcolo della chiusura e della riduzione transitiva di un grafo aciclico.

I risultati di questo lavoro vanno ben oltre l'analisi di algoritmi: vengono infatti messi in evidenza nuovi punti di contatto tra la teoria dei grafi, la teoria della probabilità, la teoria dei numeri e la teoria delle serie ipergeometriche. Il ruolo dominante che le funzioni q–ipergeometriche assumono in questo lavoro si riflette nell'uso costante della q–notazione, che si è rivelata potente mezzo per la trattazione dei nostri problemi.