Preface

Aims and scope

Big Queues aims to give a simple and elegant account of how large deviations theory can be applied to queueing problems. Large deviations theory is a collection of powerful results and general techniques for studying rare events, and has been applied to queueing problems in a variety of ways.

The strengths of large deviations theory are these: it is powerful enough that one can answer many questions which are hard to answer otherwise, and it is general enough that one can draw broad conclusions without relying on special case calculations. This latter strength has been hidden by the rather piecemeal development of the subject so far, and we feel it is now time for an account which shows that (in this case at least) abstraction can serve to simplify rather than to obscure.

We are not aiming to write an encyclopaedia on the subject, nor is this an attempt to survey the vast literature (including books by Shwartz and Weiss [91] and Chang [13]) which has evolved on this and related topics. Instead we present a certain point of view regarding the application of large deviations theory to queueing problems. Specifically, we will use the 'continuous mapping' approach, which has several benefits.

First, it suggests a style of simple heuristic argument which is easy to make rigorous.

Second, by basing our results on one key concept, the presentation is made much simpler. The continuous mapping approach lets us use exactly the same framework to describe three important scaling regimes: the large buffer regime; the regime for describing long-range dependence, which has attracted a good deal of attention in Internet traffic modelling; and the many-flows regime, which often gives better numerical approximations.

Third, this approach allows us to make very general statements about how various quantities of interest scale as the system scales, without needing to make any explicit calculations. In designing networks, it is commonly more important to understand such scaling behaviour than it is to obtain explicit answers. With the help of the continuous mapping approach, we aim to give an elementary introduction to rare-event scaling phenomena in queueing theory.

Intended readership

Big Queues targets graduate students in probability and mathematically-inclined graduate students in engineering, especially those interested in applications to communications networks. Much of the material is drawn from lecture courses given by the authors at Uppsala, Cambridge and Bangalore.

The introductory chapters and Chapter 10 on heuristics might also be of interest to the wider network-engineering research community.

Online material

The website for this book is `www.bigqueues.com`. It contains corrections, as well as an 'active bibliography' containing links to online versions of the papers cited (where available) and references to more recent articles.

Acknowledgements

Some of the material evolved from a lecture course given by N.O'C. at Uppsala University in November 1999. Acknowledgments are due to the organisers, Ingemar Kaj and Tobias Ryden, to fellow-instructors Raymond Russell and Fergal Toomey, and to the students.

Some of the material draws on a lecture course given by Stephen Turner at Cambridge. Thanks to him for making his notes available to D.J.W.

The presentation benefited from the helpful comments of Professor Anurag Kumar who invited A.J.G. to give a course on this material at the Indian Institute of Science at Bangalore.

Thanks to proof readers, including Devavrat Shah, Xiaofei Ji, Silke Meiner, and Chris Purcell.

Finally we would like to thank Venkat Anantharam, John Lewis and Frank Kelly for introducing us to large deviations theory.

Ayalvadi Ganesh, Cambridge, November 2003.
Neil O'Connell,
Damon Wischik.