Marriages Made in the Heavens: A Practical Application of Existence

DAN KALMAN
American University
Washington, DC 20016

Introduction

Existence results are familiar to mathematicians, who understand their theoretical significance. But others, including students, are sometimes perplexed by existence results. "What good is it to know something exists," they wonder, "when you have no idea how to find it?" The ultimate answer depends on some intrinsic appreciation of abstract mathematics—not all of mathematics can or should be justified on the basis of practical application. Still, abstract existence results sometimes have practical consequences, as this paper aims to demonstrate. Though the focus is on existence results, the demonstration itself is constructive: I will describe a problem I worked on in the aerospace industry that made thoroughly practical use of an existence result. Specifically, an existence theorem associated with the marriage problem for bipartite graphs was applied to a satellite communications network, matching orbiting satellites with ground stations. A briefer account, which omits most of the mathematical details, can be found in **[2]**.

The allocation problem

The problem setting is a preliminary design study for a satellite communication system. Many variables influence the design of such systems: the number of satellites, the orbits they occupy, details of the communications equipment, power requirements, etc. Normally, the complete set of satellite orbits, referred to as a constellation, is considered in total, rather than focusing on individual orbits one at a time. I was part of a team studying the effects of the constellation design on system performance.

For this preliminary design study we used highly idealized models. The earth is represented by a sphere rotating uniformly once every 24 hours about a fixed axis

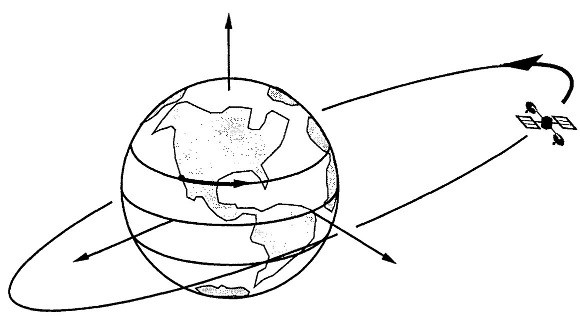

FIGURE 1
Geometrical model.

through the poles. Satellites and radio stations on the ground are points, specified in a Cartesian coordinate frame. The origin of the coordinate frame is at the center of the earth; the equator lies in the xy-plane; the axis of rotation is the z-axis. In this system, satellites travel in Keplerian orbits dictated by an inverse square force law, and we can calculate each satellite's position at any time. The ground-based radio stations move with the rotating earth, so they can also be located at any time. The operational demands on the system are assumed to be constant, with each station handling a fixed volume of message traffic in any 24-hour period.

Visibility The concept of visibility is crucial to the model. A satellite and radio station are *visible* to one another if they can communicate. In the simplest models visibility is interpreted literally, as an unobstructed line of sight. More complicated models take into account the geometry of the station and satellite, as well as physical constraints on radio transmission. These can include signal loss due to atmospheric conditions, and the sensitivity of antennas to the direction of arrival of a signal.

In these more complicated models, visibility is described in geometric terms. For example, we envision the antennas on board satellites and fixed to the ground as having a conical field of view. An arriving signal must fall within the cone to be visible. In more elaborate models the field of view can be more complicated than a simple cone, as shown in exaggerated form in FIGURE 2. To be visible from a station in the

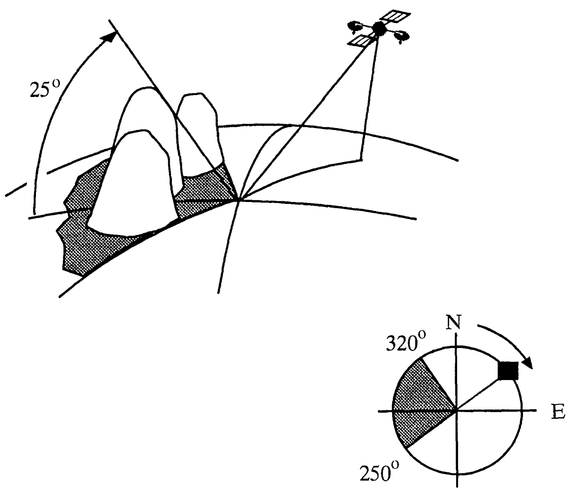

FIGURE 2
Visibility constraints.

figure, a satellite must be within the antenna's cone, but also above the mountains to the west. In this case we can define visibility as follows: Project the line joining the satellite and station onto the local horizontal plane at the station, thereby defining the satellite's heading. Measuring clockwise from due north, for headings between 250 and 320 degrees the satellite's elevation must be at least 25 degrees. For any other heading, the satellite is visible if its elevation is at least 5 degrees.

In all of these models, visibility is determined by simple geometrical relationships. At any specific time, using the instantaneous positions of the satellite and radio station, as well as the geometric constraints, one can calculate using vector analysis whether the satellite is visible to the station. To obtain an overview of the system's behavior, the calculations are repeated for many specific times, defining a *discrete time model*.

Discrete time model Although satellite and earth motions are continuous, it is reasonable to analyze them by considering a discrete set of times. Assume, for instance, that the visibility computations are performed once for each minute of the simulation. Assume, too, that a satellite can communicate for one minute with any radio station that is visible, but only with one station during that minute. On the other hand, each radio station can communicate with several satellites at once. (This asymmetry reflects the fact that we can position multiple receivers and transmitters on the ground at little expense, while resources on satellites are very limited.)

Our model also includes a predefined quota of connect-time for each radio station. These quotas are based on a projected volume of communications traffic at each station and may differ from station to station. The first station may require 90 minutes of connect-time during the simulation, the next station 30 minutes, and so on. Now we can state the fundamental problem:

> *ALLOCATION PROBLEM. At each time step, assign each satellite to one visible radio station so that, over the course of the simulation, each station achieves its quota of connect-time.*

Graph theory formulation

The allocation problem may be reformulated in terms of graph theory. To begin, consider FIGURE 3, which shows a satellite at three different times, and several radio

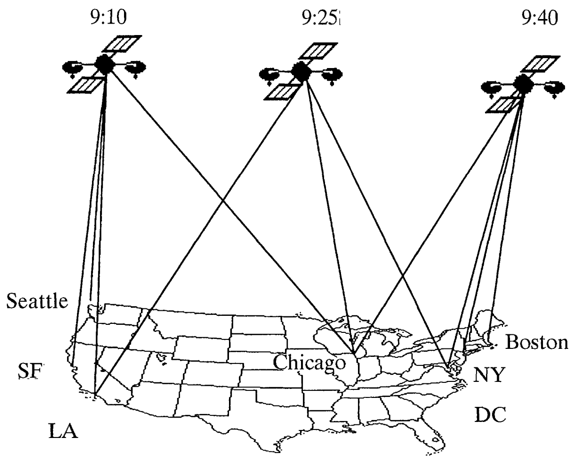

FIGURE 3
Visibility for several time steps.

stations. The lines represent visibility: at 9:10 the satellite is visible to Seattle, San Francisco, Los Angeles, and Chicago. By 9:25 the satellite has lost sight of Seattle and San Francisco, but can now see Washington, DC.

It is natural to abstract away the geographical map, leaving only a *bipartite graph* (FIGURE 4). There are two sets of vertices, or nodes: one set for the satellite at different times and the other for radio stations; each edge joins a vertex in one set to a vertex in the other.

The simple example of FIGURE 4 shows how the allocation problem is reformulated using graph theory. Define a bipartite graph V, the *visibility graph*, as follows: (i)

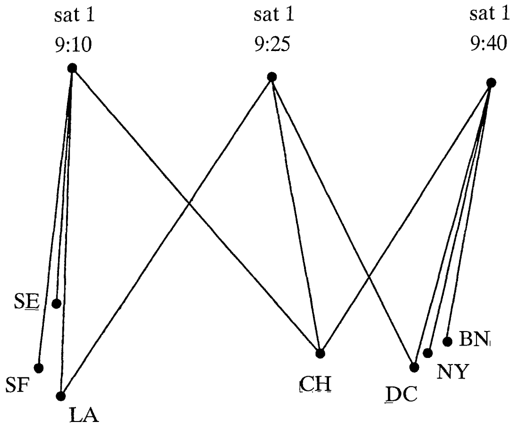

FIGURE 4
Sample visibility graph.

assign one vertex to each radio station; (ii) assign one vertex to each satellite at each of the discrete time steps; (iii) let there be an edge between the vertex for satellite s at time t and radio station r if and only if s and r are visible to one another at time t. This produces a bipartite graph, with vertices divided into two groups: satellite-time (ST)-vertices and radio station (R)-vertices.

A computer simulation is used to determine V. At each time step, positions are computed for each radio station and each satellite, and used to determine which satellites are visible to which radio stations. A typical simulation might involve 24 hours (1440 minutes), 10 satellites, and 10 radio stations. This produces over 14000 ST-vertices, only 10 R-vertices, and up to 140000 edges in the visibility graph—a result far more complicated than FIGURE 4 depicts. FIGURE 5 comes a little closer to the true situation, but it is clearly hopeless to portray anything like the true complexity of the problem.

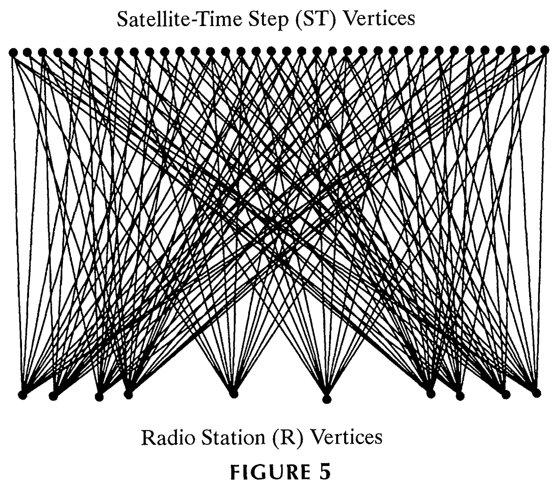

FIGURE 5
Detailed visibility graph.

Assignment subgraphs Part of the allocation problem is to assign each satellite at each time step to just *one* radio station. In the graph-theoretic formulation, that means selecting just one edge emanating from each ST-vertex, thus defining a subgraph, which we call an *assignment subgraph.* The bold lines in FIGURE 6 show one possible assignment subgraph for the graph of FIGURE 5.

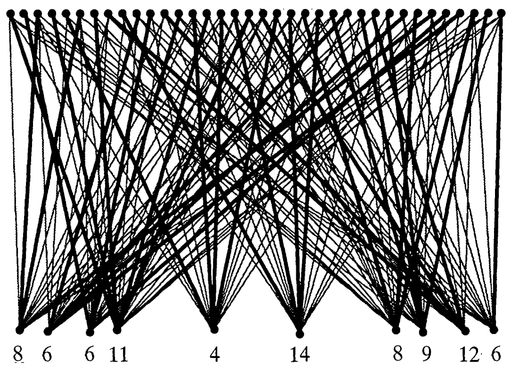

FIGURE 6
Assignment subgraph.

The number of edges in a graph G incident at a vertex v is called the *degree* of v in G, and denoted $\deg(v, G)$. An assignment subgraph A of the visibility graph V is thus characterized by the requirement that $\deg(st, A) \leq 1$ for every ST-vertex st.

The assignment problem has another requirement: each radio station must be connected to satellites for a predetermined quota of minutes. In the assignment subgraph, each edge represents one minute of connect-time. Thus the connect-time quota for a radio station r dictates a minimum number of incident edges (i.e., the degree) at r in the assignment subgraph.

Connect time quotas are shown for each R-vertex in FIGURE 6. The assignment subgraph shown clearly fails to satisfy the allocation problem because, for example, the first R-vertex has 3 incident bold edges—less than the quota of 8.

We can now state the allocation problem in graph-theoretic terms.

ALLOCATION GRAPH PROBLEM: *Given a visibility graph V and a connect-time quota $q(r)$ for each R-vertex, find an assignment subgraph A such that* $\deg(r, A) \geq q(r)$ *for every R-vertex r.*

It is not obvious at the outset whether this kind of problem is solvable. Next we will develop some necessary conditions for solvability. These conditions will lead to an existence theorem—our desired existence result—that characterizes the solvability of allocation graph problems.

Necessary conditions

It is easy to see that the sample assignment subgraph in FIGURE 6 does not solve the allocation problem, but it is less obvious whether *any* solution exists. A closer look at FIGURE 6 reveals that each R-vertex has far too few incident bold lines. This suggests that, no matter how we rearrange the assignments, we will be unable to fill the quota at each R-vertex. This is indeed so, and there is a simple proof. Do you see it?

Because at most one edge in the assignment subgraph can meet each ST-vertex, the total number of edges cannot exceed the number of ST-vertices, 36. The allocation problem requires enough bold edges to match the number beneath each R-vertex. Since the sum of those numbers is much greater than 36, the number of bold edges available, this allocation problem is unsolvable.

The actual simulations that I worked with had much more complicated graphs than that in FIGURE 6, but the same reasoning applies. If the total number of ST-vertices is less than the sum of all the connect-time quotas, then the allocation problem will be unsolvable. This gives a necessary condition for solvability: The sum of all the connect-time quotas $q(r)$ cannot exceed the number of ST-vertices.

Other necessary conditions arise just as naturally. The graph in FIGURE 7, for instance, admits no solution to the allocation problem, because only one edge in V is

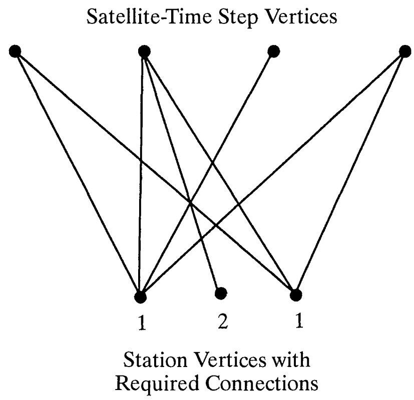

FIGURE 7
Required degree too high at one station.

incident at the second R-vertex, but the allocation problem requires two edges there. This suggests a second necessary condition for the allocation problem to be solvable: $q(r) \leq \deg(r, V)$ for every R-vertex r.

FIGURE 8 illustrates yet another unsolvable allocation problem. As before, a general principle is at work, but this time a little more subtly. To see how it works, consider

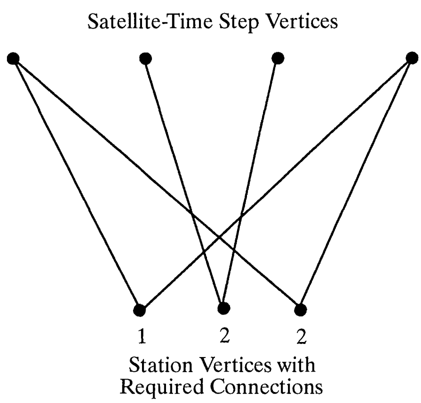

FIGURE 8
Total required degrees too high for a set of stations.

only the part of the graph that involves the first and last R-vertices. If we ignore the middle R-vertex, we may as well ignore the two middle ST-vertices as well. What remains are two ST-vertices, two R-vertices, and 4 edges. The allocation problem is clearly unsolvable in this subgraph, for the sum of the required degrees (3) exceeds the number of ST nodes (2). Since the allocation problem cannot be solved for this subgraph, it cannot be solved for the original graph either.

A unifying principle The preceding example illustrates a more general principle. For any subset E of the R-vertices, consider the subgraph V_E consisting of the edges that touch elements of E, together with the endpoints of these edges. Let $N_{ST}(V_E)$ be the number of ST-vertices in V_E. For the original allocation problem to be solvable, we must have, for every non-empty subset E of R-vertices,

$$\sum_{r \in E} q(r) \leq N_{ST}(V_E). \tag{1}$$

Note that this one principle subsumes both of the preceding examples. If E is the full set of R-vertices, then $V_E = V$, and the corresponding necessary condition is what was presented in the first example. If $E = \{r\}$, for any single R-vertex, then $N_{ST}(V_E) = \deg(r, V)$, and the necessary condition is as in the second example.

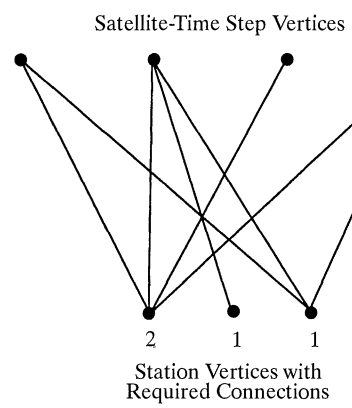

FIGURE 9
Required degrees feasible.

FIGURE 9 shows a graph in which the allocation problem *is* solvable. It is easy to check that the condition (1) holds for every possible E.

The existence theorem

The necessity of the solution condition (1) should now be intuitively clear. What is less clear, but still true, is that the solution condition is also sufficient: If (1) holds for every nonempty set E of R-vertices, then a solution to the allocation problem exists. This result is equivalent to Hall's theorem,[1] a classical result in graph theory. Hall's

[1]According to [3], Philip Hall deduced the result as a theorem in set theory in 1935.

theorem is often cited in the context of the so-called *marriage problem*; in [**4**, p. 159], the result is labeled *Hall's Marriage Theorem.*

The marriage problem is a special case of the allocation problem, with $q(r) = 1$ for every r. In its traditional formulation, the R-vertices represent maidens, the ST-vertices represent bachelors, and an edge indicates that a particular bachelor and maiden are acquainted with each other. The problem is for each bachelor to propose marriage to just one maiden in such a way that each maiden receives at least one proposal. It is assumed that there are just as many bachelors as maidens. Hall's theorem asserts that a solution exists if and only if every set of k maidens is connected to a set of at least k bachelors.

This condition is clearly necessary: if any k maidens are connected to fewer than k bachelors, there will not be enough bachelors to go around. Sufficiency may be proved by induction on the number of bachelors (and maidens). Briefly, the induction step can be handled by considering two cases. In the first case, each set of k maidens is connected to a set of at least $k+1$ bachelors. In this case, we match one bachelor to one maiden, and the induction hypothesis implies that everyone else can also be matched up. In the second case, some set of k maidens is connected to exactly k bachelors. In this case we first argue (by induction) that these k maidens and bachelors can be paired, and then argue (again by induction) that the remaining matches can be made as well.

The satellite allocation problem can be reduced to a marriage problem as follows. Create a new graph V_M by creating $q(r)$ duplicate vertices for each R-vertex r; each of these $q(r)$ vertices is connected by an edge to each ST-vertex that r was connected to in the original graph. If the marriage problem can be solved in V_M, then we will have $q(r)$ ST-vertices matched with the duplicate vertices for r; this specifies how to assign ST-vertices to r in V.

We must also assure that the number of R-vertices equals the number of ST-vertices in V_M. Observe that if there are too few ST-vertices, the original problem is not solvable. If there are too many ST-vertices, we can simply add enough R-vertices to make the two sets compatible, and assume that each of these new R-vertices is connected to every ST-vertex.

Now we see that the condition defined by (1) translates into the necessary and sufficient condition for solvability of the marriage problem. A solution to the marriage problem, however, induces a solution to the satellite allocation problem. It may happen that in formulating the marriage problem, some extra R-vertices were added. In this case, the matches of these extra R-vertices with ST-vertices in the solution of the marriage problem will be discarded in translating back to the satellite allocation problem. But the result will still be an assignment subgraph that meets all the connection-time quotas.

With (1) established as a necessary and sufficient condition for solvability of the allocation graph problem, deciding whether a satellite constellation is capable of meeting its performance objectives is reduced to a computation. Note that this computation does not provide any practical operational guidance: we still do not know how to assign the satellites to ground stations. The existence of a solution has a practical significance in a different direction, providing a metric of system performance. This idea is developed next.

A practical use for existence

The allocation problem arose in the consideration of various satellite constellations for a communications system. The existence result provided a way to compare different constellations. Beyond the simple observation that for some constellations the allocation problem is solvable and for others it is not, the existence result led to the determination of an optimal data transmission rate for each constellation. Here is how it worked.

First, the calculations necessary to check the conditions for solvability were incorporated into the computer simulation. With N R-vertices, there are $2^N - 1$ nontrivial choices for E; for each, it is necessary to tabulate the number of ST-vertices connected to E. Although the dependence on N is exponential, N was small enough in our problems to permit a direct calculation in reasonable time.

There are well-known algorithms for solving graph matching problems by reducing them to flow optimization problems.[2] However, for the visibility graphs encountered in the satellite design problem, these algorithms require more calculations (by several orders of magnitude) than simply checking the necessary and sufficient conditions for solvability. Thus, it was feasible to compute whether a solution existed—but not to find a solution.

Second, to the degree of accuracy of the model, it is reasonable to assume that the values $q(r)$ are inversely proportional to the rate of data transmission. For example, doubling the transmission rate should halve the amount of connect time required.

As a general rule, faster data transmission is more expensive. This is certainly so for computer modems, and it applies even more stringently to satellites. The intrinsic cost of faster data rates is compounded by increased power requirements, which generally translate into greater weight and complexity of the satellite. So it is of interest to estimate the minimal feasible data rate for a satellite constellation, defined as the lowest data rate for which the assignment problem is solvable.

Note here that changing the data rate has no effect on the visibility graph—it simply increases or reduces the values of the $q(r)$. If a very high data rate is set, the values of the $q(r)$ will be low, and it is likely that the allocation problem can be solved. With a low data rate, the allocation problem is harder to solve.

Here, then, is how to estimate the minimal feasible data rate. Run the simulation once to compute V. Select an initial choice for the data rate and, using the existence result, determine whether the allocation problem is solvable. If it is, lower the transmission rate; otherwise, raise the transmission rate. Then check the conditions for solvability again. Repeating this process a few times will usually establish the minimal feasible transmission rate within a few percent.

Conclusion

Ultimately, the minimal feasible transmission rate became just one of many criteria that were used to compare competing satellite system designs. I spent considerable time running the simulation and computing the optimal transmission rate for many satellite constellations, and that contributed to a much larger tradeoff analysis. In the process, solutions to satellite allocation problems were neither desired nor obtained. Much later in the design process, algorithms would have to be developed to actually assign the satellites to communicate with particular radio stations at particular times, and in all likelihood, these assignments would not be optimal. But at the point of the

[2] See [**1**, **4**]; the first of these includes an analysis of the computational complexity of the algorithms.

design process I have been describing, that level of detail was not required. Rather, the theoretical solvability of the optimal allocation problem was used to determine one measure of system performance. And that is how an existence result was used in a completely practical setting.

REFERENCES

1. Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms*, MIT Press, Cambridge, MA, 1990, 587–599.
2. Dan Kalman. When existence is enough, *Math Horizons*, April 1997, 23–26.
3. Oystein Ore and Robin J. Wilson. *Graphs and Their Uses*, New Mathematical Library, Vol. 34, Mathematical Association of America, Washington, DC, 1990, 53–59.
4. Alan Tucker. *Applied Combinatorics*, John Wiley & Sons, New York, NY, 1995, 137–167.

REVIEWS (continued from page 157)

Shulman, Polly, From Muhammad Ali to Grandma Rose, *Discover* 19 (12) (December 1998) 85–89; `http://www.discover.com/decissue/smallworld.html` . Peterson, Ivars, Close connections: It's a small world of crickets, nerve cells, computers, and people, *Science News* 154 (22 August 1998) 124–126. Harris, John M., and Michael J. Mossinghoff, The eccentricities of actors, *Math Horizons* (February 1998) 23–25. Strogatz, Steven H., and Duncan J. Watts, Collective dynamics of "small-world" networks, *Nature* 393 (4 June 1998) 440–442.

Modern folklore claims that everyone on the planet is connected to everyone else through a relatively short chain of acquaintances. This claim was investigated by psychologist Stanley Milgram in the 1960s, concretized in the play and film *Six Degrees of Separation* by John Guare, and popularized in a game of trying to trace connections of actors to the actor Kevin Bacon through joint appearances in films. Mathematicians are familiar with a similar phenomenon, the *Erdős number* of a mathematician. We are talking eccentricity here; in graph theory, the maximum distance from a particular vertex to any other vertex in a graph is the *eccentricity* of the vertex. Harris and Mossinghoff document that the eccentricity of Kevin Bacon in the "Hollywood" graph is 7, which is the minimum eccentricity of any actor, putting Bacon into what is known as the *center* of the graph. Such networks are neither regular (every node has the same small number of links to neighboring points) nor sparse (few connections relative to the number of nodes). Strogatz and Watts showed that introducing a few random connections into a regular graph can greatly decrease the average path length between two nodes. Graphs with a small average path length they call *small-world networks*, and they cite as examples the neural network of the worm *C. elegans*, the power grid of the western U.S., and the Hollywood graph. Small-world networks are important in the spread of disease, the diffusion of trade goods, and the transmission of information (including marketing over the Internet), as Peterson notes.

Petković, Miodrag, *Mathematics and Chess: 110 Entertaining Problems and Solutions*, Dover, 1998; v + 133 pp, $5.95 (P). ISBN 0-486-29432-3.

This book uses chess and the chessboard as occasions for mostly mathematical puzzles, though some chess puzzles occur too. Naturally, rook polynomials occur, as do knight's tours and their generalizations, plus domino coverings, dissections, and generalized chessboards. The puzzles are fun, solutions are provided, and the reader will see much mathematics applied. There are a few references but only to specific results, rather than to further reading; and lamentably there is no index.