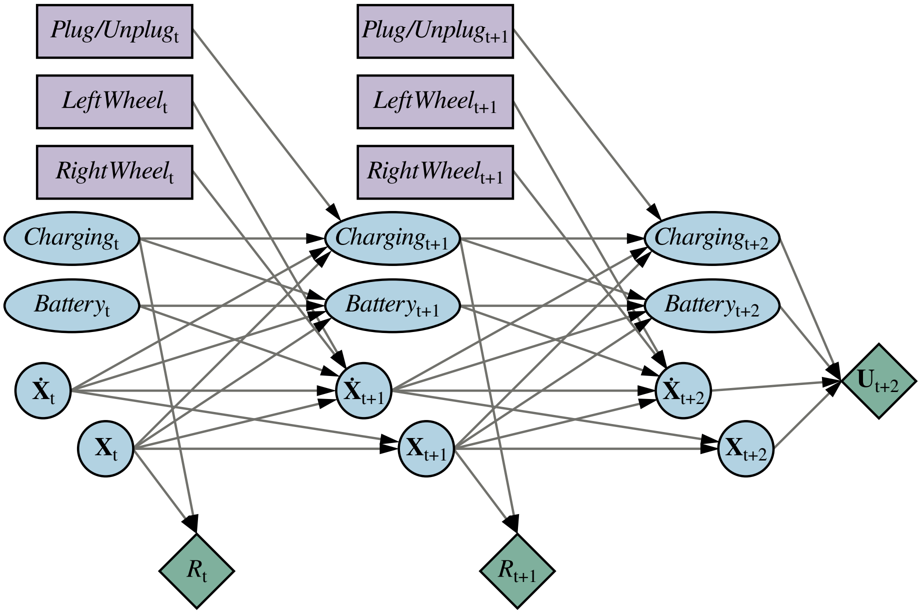

Figure 17.4 A dynamic decision network for a mobile robot with state variables for battery level, charging status, location, and velocity, and action variables for the left and right wheel motors and for charging.

Chapter 2; they typically have an exponential complexity advantage over atomic representations and can model quite substantial real-world problems.

Figure 17.4, which is based on the DBN in Figure 14.13(b) (page 486), shows some elements of a slightly realistic model for a mobile robot that can charge itself. The state S_t is decomposed into four state variables:

- $\mathbf{X}_t$ consists of the two-dimensional location on a grid plus the orientation;
- $\dot{\mathbf{X}}_t$ is the rate of change of $\mathbf{X}_t$;
- $Charging_t$ is true when the robot is plugged in to a power source;
- $Battery_t$ is the battery level, which we model as an integer in the range $0,\ldots,5$.

The state space for the MDP is the Cartesian product of the ranges of these four variables. The action is now a set $\mathbf{A}_t$ of action variables, comprised of $Plug/Unplug$, which has three values (*plug*, *unplug*, and *noop*); *LeftWheel* for the power sent to the left wheel; and *RightWheel* for the power sent to the right wheel. The set of actions for the MDP is the Cartesian product of the ranges of these three variables. Notice that each action variable affects only a subset of the state variables.

The overall transition model is the conditional distribution $\mathbf{P}(\mathbf{X}_{t+1}|\mathbf{X}_t,\mathbf{A}_t)$, which can be computed as a product of conditional probabilities from the DDN. The reward here is a single variable that depends only on the location $\mathbf{X}$ (for, say, arriving at a destination) and *Charging*, as the robot has to pay for electricity used; in this particular model, the reward doesn't depend on the action or the outcome state.

The network in Figure 17.4 has been projected two steps into the future. Notice that the network includes nodes for the *rewards* for times t and $t+1$, but the *utility* for time $t+2$. This