Here $B(b, z, u)$ is computed using the Bayes filter. In robotics, we only have a choice over the control action u, we cannot pick z. Consequently, we consider the conditional entropy of the control u, with the measurement integrated out:

$$
\begin{aligned}
H_b(x' \mid u) &\approx E_z[H_b(x' \mid z, u)] \\
&= \int\int H_b(x' \mid z, u)\; p(z \mid x')\; p(x' \mid u, x)\; b(x)\; dz\; dx'\; dx
\end{aligned}
\tag{17.3}
$$

INFORMATION GAIN

Notice that this is only an approximation, as the final expression inverts the oder of a summation and a logarithm. The *information gain* associated with action u in belief b is thus given by the difference

$$
I_b(u) = H_p(x) - E_z[H_b(x' \mid z, u)] \tag{17.4}
$$

17.2.2 Greedy Techniques

The expected information gain lets us phrase the exploration problem as a decision-theoretic problem of the type addressed in previous chapters. In particular, let $r(x, u)$ be the cost of applying control action u in state x; here we assume $r(x, u) < 0$ to keep our notation consistent. Then the optimal greedy exploration for the belief b maximizes the difference between the information gain and the costs, weighted by a factor α.

$$
\pi(b) = \underset{u}{\operatorname{argmax}}\; \alpha\; \underbrace{(H_p(x) - E_z[H_b(x' \mid z, u)])}_{\text{expected information gain}} + \underbrace{\int r(x, u)\; b(x)\; dx}_{\text{expected costs}} \tag{17.5}
$$

The factor α relates information to the cost of executing u. It specifies the value a robot assigns to information, which measures the price it is willing to pay in terms of costs for obtaining information.

Equation (17.5) resolves to

$$
\begin{aligned}
\pi(b) &= \underset{u}{\operatorname{argmax}}\; -\alpha\; E_z[H_b(x' \mid z, u)] + \int r(x, u)\; b(x)\; dx \\
&= \underset{u}{\operatorname{argmax}} \int [r(x, u) - \alpha \int\int H_b(x' \mid z, u)\; p(z \mid x') \\
&\qquad p(x' \mid u, x)\; dz\; dx']\; b(x)\; dx
\end{aligned}
\tag{17.6}
$$

In short, to understand the utility of the control u, we need to compute the expected entropy after executing u and observing. This expected entropy is obtained by integrating over all possible measurements z that we might be