2.3 Show that in the case of two actions, the softmax operation using the Gibbs distribution becomes the logistic, or sigmoid, function commonly used in artificial neural networks. What effect does the temperature parameter have on the function?

SOLUTION The Gibbs distribution for a system with two possible actions is given by:

$$P(a_1) = \frac{e^{-Q(a_1)/\tau}}{e^{-Q(a_1)/\tau} + e^{-Q(a_2)/\tau}}$$

$$P(a_2) = \frac{e^{-Q(a_2)/\tau}}{e^{-Q(a_1)/\tau} + e^{-Q(a_2)/\tau}}$$

where $Q(a_1)$ and $Q(a_2)$ are the values associated with the two actions, and τ is the temperature parameter.

Now let's define the logistic (sigmoid) function, which is commonly used in artificial neural networks. The sigmoid function is given by:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

To show that the softmax operation becomes the logistic (sigmoid) function in the case of two actions, let's express $P(a_1)$ in terms of the sigmoid function:

$$P(a_1) = \frac{e^{Q(a_1)/\tau}}{e^{Q(a_1)/\tau} + e^{Q(a_2)/\tau}} = \frac{1}{1 + e^{(Q(a_2)-Q(a_1))/\tau}}$$

Comparing this with the sigmoid function definition, we can see that $P(a_1)$ is a sigmoid function with input $-(Q(a_1) - Q(a_2)/\tau$:

$$P(a_1) = \sigma((-(Q(a_1) - Q(a_2))/\tau)$$

Thus, we have shown that in the case of two actions, the softmax operation using the Gibbs distribution becomes the logistic (sigmoid) function commonly used in artificial neural networks.

The temperature parameter τ has an effect on the softmax function as follows:

1. $\tau \to 0$, the softmax function becomes more "deterministic." The highest-valued action approaches a probability of 1, while the others approach 0.

2. When $\tau \to \infty$, the softmax function becomes more "uniform." The probabilities for all actions converge to $1/n$, where n is the number of actions. In the case of two actions, this means that the probabilities for both actions approach 0.5.

3. For intermediate values of τ, the softmax function smoothly interpolates between the deterministic and uniform cases, producing a balance between exploration and exploitation.

In the context of the sigmoid function, the temperature parameter τ affects the steepness of the curve. Higher values of τ result in a flatter curve (smoother transition between 0 and 1), while lower values of τ result in a steeper curve (sharper transition between 0 and 1).