Let $X$ and $Y$ be countable sets. We denote $n$-tuples $(x_1, \ldots, x_n) \in X^n$ and $(y_1, \ldots, y_n) \in Y^n$ by $x^n$ and $y^n$, respectively. For $k = 1, 2, \ldots$, let $X_k$ and $Y_k$ be $X$- and $Y$-valued random variables, respectively. Then for $n \geq 2$, we can always write

$$\mathbb{P}(Y^n = y^n, X^n = x^n) = \mathbb{P}(Y_n = y_n | Y^{n-1} = y^{n-1}, X^n = x^n) \cdot \mathbb{P}(X_n = x_n | Y^{n-1} = y^{n-1}, X^{n-1} = x^{n-1}) \cdot \mathbb{P}(Y^{n-1} = y^{n-1}, X^{n-1} = x^{n-1}).$$

Conversely, let $a_1(\cdot)$ be a probability mass function (PMF) on $X$, and for each $x \in X$, let $b_1(\cdot | x)$ be a PMF on $Y$. Similarly, for $n \geq 2$, let $a_n(\cdot | y^{n-1}, x^n)$ be a PMF on $X$, and let $b_n(\cdot | y^{n-1}, x^n)$ be a PMF on $Y$. Then a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variables $X_k$ and $Y_k$ exist$^1$ such that

$$\mathbb{P}(X_1 = x) = a_1(x) \quad \text{and} \quad \mathbb{P}(Y_1 = y | X_1 = x) = b_1(y | x),$$

and for $n \geq 2$,

$$\mathbb{P}(Y_n = y_n | Y^{n-1} = y^{n-1}, X^n = x^n) = b_n(y_n | y^{n-1}, x^n) \quad \mathbb{P}(X_n = x_n | Y^{n-1} = y^{n-1}, X^{n-1} = x^{n-1}) = a_n(x_n | y^{n-1}, x^{n-1}).$$

One is often interested in computing quantities such as $\mathbb{P}(X_n = x_n | Y^n = y^n)$. To this end, let

$$r_n(x_n, y^n) := \mathbb{P}(X_n = x_n, Y^n = y^n),$$

and note that

$$\mathbb{P}(X_n = x_n | Y^n = y^n) = r_n(x_n, y^n) \sum_{x \in X} r_n(x, y^n).$$

Clearly, $r_1(x, y) = a_1(x) b_1(y | x)$. Now, suppose that for $n \geq 2$, $b_n(y_n | y^{n-1}, x^n) = \beta_n(y_n | y^{n-1}, x_n)$ and $a_n(x_n | y^{n-1}, x^{n-1}) = \alpha_n(x_n | y^{n-1}, x_{n-1})$ for some functions $\beta_n$ and $\alpha_n$. Under these Markov-type assumptions, it is a trivial calculation to show that

$$r_n(x_n, y^n) = \beta_n(y_n | y^{n-1}, x_n) \sum_{x \in X} \alpha_n(x_n | y^{n-1}, x) r_{n-1}(x, y^{n-1}), \quad n \geq 2.$$

$^1$If the sequence of random variables is not finite, we need to apply to Kolmogorov’s Theorem.