Extract from MS Thesis of Grzegorz Wolny

Grzegorz Wolny

[1] Warsaw University, Banacha 2, 02-097 Warsaw, Poland
[2] Telekomunikacja Polska, Research and Development Centre Branch, Obrzeżna 7, 02-691 Warsaw, Poland gw201113@students.mimuw.edu.pl

Abstract. This document contains the proofs from chapter 4 of the paper "Time distribution in Structural Workflow Nets".

1 Cumulative Distribution Function

Let's compute the cumulative distribution function for random variable X with density being an exponential polynomial:

$$g(x) = \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} a_{i,j}\, x^i\, e^{-\alpha_{i,j}\, x}\, \mathbb{1}_{[0,\infty)}(x)$$

The cumulative distribution function of the random variable X is given by:

$$F(t) = \int_{-\infty}^{t} g(x)\,\mathrm{d}x = \int_{-\infty}^{t} \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} a_{i,j} x^i e^{-\alpha_{i,j} x} \mathbb{1}_{[0,\infty)}(x)\,\mathrm{d}x$$

For $t \leq 0$ we have $F(t) = 0$. Further calculations are made with the assumption that $t > 0$.

$$\begin{aligned}
F(t) &= \int_{0}^{t} \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} a_{i,j}\, x^i\, e^{-\alpha_{i,j}\, x}\,\mathrm{d}x = \\
&= \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \int_{0}^{t} a_{i,j}\, x^i\, e^{-\alpha_{i,j}\, x}\,\mathrm{d}x = \\
&= \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}} \left(\Gamma\left(i+1\right) - \Gamma\left(i+1, \alpha_{i,j} t\right)\right) = \\
&= \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, \Gamma\left(i+1\right) - \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, \Gamma\left(i+1, \alpha_{i,j} t\right) = \\
&= \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, i! - \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, i!\, e^{-\alpha_{i,j}\, t} \sum_{k=0}^{i} \frac{\left(\alpha_{i,j} t\right)^k}{k!} = \\
&= \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, i! - \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \sum_{k=0}^{i} \frac{a_{i,j}}{\alpha_{i,j}^{i-k+1}} \frac{i!}{k!}\, t^k\, e^{-\alpha_{i,j}\, t}
\end{aligned}$$

The value of derivative $\int_0^t a_{i,j}\, x^i\, e^{-\alpha_{i,j}\, x}\, \mathrm{d}x$ was computed using the *Mathematica* software. Function Γ is defined by:

$$\Gamma(n) = (n-1)! \quad n \in \mathbb{N}^+$$
$$\Gamma(n,x) = (n-1)!\, e^{-x} \sum_{k=0}^{n-1} \frac{x^k}{k!} \quad n \in \mathbb{N}^+ \tag{1}$$

Finally, the cumulative distribution function of the random variable X with density $g(x)$ is given by:

$$F(t) = \left(\sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, i! - \sum_{i=0}^{\max} \sum_{j=1}^{\mathrm{pow}_i} \sum_{k=0}^{i} \frac{a_{i,j}}{\alpha_{i,j}^{i-k+1}} \frac{i!}{k!}\, t^k\, e^{-\alpha_{i,j}\, t} \right) \mathbb{1}_{[0,\infty)}(t) \tag{2}$$

2 Sum

Consider random variable being the sum of n independent random variables $X_1, \ldots, X_n$. We will denote it by $SUM(X_1, \ldots, X_n)$, its density by $s_{X_1,\ldots,X_n}$, and its cdf by $S_{X_1,\ldots,X_n}$.

Proposition 1. *For the sum of independent random variables $X_1, \ldots, X_n$ we have:*

$$SUM(X_1, \ldots, X_n) = SUM(X_{\delta(1)}, \ldots, X_{\delta(n)}) \quad \delta \in S_n \tag{3}$$

$$SUM(X_1, \ldots, X_n)(t) = SUM(SUM(X_1, \ldots, X_{n-1}), X_n) \tag{4}$$

Proof. These equalities follow directly from associativity and commutativity of the sum operation.

From proposition 1 we deduce that in order to compute the density of sum of n independent random variables it is sufficient to know how to compute the density of sum of two such variables.

The following proposition is well known:

Proposition 2. *The density of sum of independent continuous random variables is random variable with density being the convolution of densities of arguments.*

The easy conclusion from this proposition is:

Proposition 3. *The density of sum of two independent random variables X and Y with densities f_X and f_Y, such that $f_X(x) = 0$ and $f_Y(x) = 0$ for $x < 0$ and continuous for $x \geq 0$ is given by:*

$$f_{X+Y}(t) = \begin{cases} \int_0^t f_X(x) f_Y(t-x)\, \mathrm{d}x & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{5}$$

Let's compute the density of sum of two independent random variables X and Y with densities $f(x)$ and $g(x)$ being exponential polynomials:

$$f(x) = \sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} a_{i,j}\, x^i\, e^{-\alpha_{i,j}\, x} \mathbb{1}_{[0,\infty)}(x)$$

$$g(x) = \sum_{i=0}^{\mathrm{mY}} \sum_{j=1}^{\mathrm{pY}_i} b_{i,j}\, x^i\, e^{-\beta_{i,j}\, x} \mathbb{1}_{[0,\infty)}(x)$$

We will obtain the result which is also the exponential polynomial. Densities f and g fulfill the assumptions of proposition 3. For $t < 0$ $s_{X,Y}(t) = 0$. For $t \geq 0$ we have:

$$\begin{aligned}
s_{X,Y}(t) &= \int_0^t f(x) g(t-x)\, \mathrm{d}x = \\
&= \int_0^t \left(\sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} a_{i,j}\, x^i\, e^{-\alpha_{i,j}\, x} \right) \sum_{i=0}^{\mathrm{mY}} \sum_{j=1}^{\mathrm{pY}_i} b_{i,j}\, (t-x)^i\, e^{-\beta_{i,j}\, (t-x)}\, \mathrm{d}x = \\
&= \int_0^t \left(\sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} a_{i,j}\, x^i\, e^{-\alpha_{i,j}\, x} \right) \sum_{k=0}^{\mathrm{mY}} \sum_{l=1}^{\mathrm{pY}_k} b_{k,l}\, (t-x)^k\, e^{-\beta_{k,l}\, (t-x)}\, \mathrm{d}x = \\
&= \sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} \sum_{k=0}^{\mathrm{mY}} \sum_{l=1}^{\mathrm{pY}_k} a_{i,j}\, b_{k,l} \int_0^t x^i (t-x)^k\, e^{-\alpha_{i,j}\, x}\, e^{-\beta_{k,l}\, (t-x)}\, \mathrm{d}x = \\
&= \sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} \sum_{k=0}^{\mathrm{mY}} \sum_{l=1}^{\mathrm{pY}_k} a_{i,j}\, b_{k,l}\, e^{-\beta_{k,l}\, t} \int_0^t x^i (t-x)^k\, e^{-(\alpha_{i,j} - \beta_{k,l})\, x}\, \mathrm{d}x
\end{aligned}$$

We will consider two cases. In the first case ($\alpha_{i,j} = \beta_{k,l}$), we add to the result the following component (the derivative was computed using the *Mathematica* software):

$$I_1 = a_{i,j}\, b_{k,l}\, e^{-\beta_{k,l} t} \int_0^t x^i (t-x)^k\, \mathrm{d}x = a_{i,j}\, b_{k,l} \frac{i!\, k!}{(i+k+1)!}\, t^{i+k+1}\, e^{-\beta_{k,l}\, t} \quad (6)$$

Observe, that we can rewrite this component as:

$$I_1 = a_{i,j}\, b_{k,l} \frac{i!\, k!}{(i+k+1)!}\, t^{i+k+1}\, e^{-\alpha_{i,j}\, t}$$

In the second case, when $\alpha_{i,j} \neq \beta_{k,l}$, we add to the result the following component:

$$\begin{aligned}
I_2 &= a_{i,j}\, b_{k,l}\, e^{-\beta_{k,l} t} \int_0^t x^i (t-x)^k\, e^{-(\alpha_{i,j}-\beta_{k,l})\, x}\, \mathrm{d}x = \\
&= a_{i,j}\, b_{k,l}\, e^{-\beta_{k,l} t} \int_0^t x^i (\sum_{d=0}^{k} (-1)^{k-d} \binom{k}{d} t^d\, x^{k-d})\, e^{-(\alpha_{i,j}-\beta_{k,l})\, x}\, \mathrm{d}x = \\
&= \sum_{d=0}^{k} (-1)^{k-d} \binom{k}{d} a_{i,j}\, b_{k,l}\, t^d\, e^{-\beta_{k,l}\, t} \int_0^t x^{i+k-d} e^{-(\alpha_{i,j}-\beta_{k,l})\, x}\, \mathrm{d}x
\end{aligned}$$

This derivative was also computed using the *Mathematica* software and is equal to:

$$\int_0^t x^s\, e^{-\alpha x}\, \mathrm{d}x = (s! - \Gamma\,(s+1, \alpha t))\, \alpha^{-s+1},$$

where function Γ is defined by equation 1. So we have:

$$\begin{aligned}
I_2 = &\sum_{d=0}^{k} (-1)^{k-d} \binom{k}{d} (i+k-d)! \frac{a_{i,j}\, b_{k,l}}{(\alpha_{i,j}-\beta_{k,l})^{i+k-d+1}}\, t^d\, e^{-\beta_{k,l}\, t} + \\
&- \sum_{d=0}^{k} (-1)^{k-d} \binom{k}{d} \frac{a_{i,j}\, b_{k,l}}{(\alpha_{i,j}-\beta_{k,l})^{i+k-d+1}}\, t^d\, e^{-\beta_{k,l}\, t}\, \Gamma\,(i+k-d+1, t\,(\alpha_{i,j}-\beta_{k,l}))
\end{aligned}$$

Let's represent component I_2, so that $I_2 = I_{2,1} - I_{2,2}$, where:

$$I_{2,1} = \sum_{d=0}^{k} (-1)^{k-d} \binom{k}{d} (i+k-d)! \frac{a_{i,j}\, b_{k,l}}{(\alpha_{i,j}-\beta_{k,l})^{i+k-d+1}}\, t^d\, e^{-\beta_{k,l}\, t} \tag{7}$$

and:

$$\begin{aligned}
I_{2,2} = &\sum_{d=0}^{k} (-1)^{k-d} \binom{k}{d} \frac{a_{i,j}\, b_{k,l}}{(\alpha_{i,j}-\beta_{k,l})^{i+k-d+1}}\, t^d\, e^{-\beta_{k,l}\, t} \cdot \\
&\cdot \left((i+k-d)!\, e^{-(\alpha_{i,j}-\beta_{k,l})\, t} \sum_{c=0}^{i+k-d} \frac{t^c\, (\alpha_{i,j}-\beta_{k,l})^c}{c!} \right) = \\
= &\sum_{d=0}^{k} \sum_{c=0}^{i+k-d} (-1)^{k-d} \binom{k}{d} \frac{(i+k-d)!}{c!} \frac{a_{i,j} b_{k,l}}{(\alpha_{i,j}-\beta_{k,l})^{i+k-d-c+1}}\, t^{d+c}\, e^{-\alpha_{i,j}\, t}
\end{aligned} \tag{8}$$

The density function being sum of components of the form I_1, $I_{2,1}$ and $I_{2,2}$, has the desired form. We will simplify it. Observe that the proposition 3 allows us to write $s_{X,Y}(t) = s_{Y,X}(t)$. Observe also that all exponents of e in formulas of the form $I_{2,1}$ occur in the formula for the density of Y, and all exponents of e in the formula $I_{2,2}$ occur in the density of X. Observe that since $s_{X,Y}(t) = s_{Y,X}(t)$, the components of the form I_1 will be reduced (probably using the last observation).

The remained components of the form $I_{2,1}$, $I_{2,2}$ and $I'_{2,1}$, $I'_{2,2}$ (analogous formulas from density of $s_{Y,X}(t)$) must be equal. So we can conclude that the sum of components of the form $I_{2,2}$ can be replaced by the sum of components of the form $I'_{2,1}$ (which are simpler). As the final result we obtain the following formula for the density of $s_{X,Y}$:

$$s_{X,Y}(t) = \sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} \sum_{k=0}^{\mathrm{mY}} \sum_{l=1}^{\mathrm{pY}_k} \frac{i!\, k!\, a_{i,j}\, b_{k,l}}{(i+k+1)!} t^{i+k+1}\, e^{-\alpha_{i,j} t}\, \mathbb{1}_{\alpha_{i,j}=\beta_{k,l}} +$$
$$+ \sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} \sum_{d=0}^{i} \sum_{k=0}^{\mathrm{mY}} \sum_{l=1}^{\mathrm{pY}_k} \binom{i}{d} \frac{(-1)^{i-d}\,(k+i-d)!\, a_{i,j}\, b_{k,l}}{(\beta_{k,l}-\alpha_{i,j})^{k+i-d+1}} t^d\, e^{-\alpha_{i,j} t}\, \mathbb{1}_{\alpha_{i,j}\neq\beta_{k,l}} +$$
$$+ \sum_{k=0}^{\mathrm{mY}} \sum_{l=1}^{\mathrm{pY}_k} \sum_{d=0}^{k} \sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} \binom{k}{d} \frac{(-1)^{k-d}\,(i+k-d)!\, a_{i,j}\, b_{k,l}}{(\alpha_{i,j}-\beta_{k,l})^{i+k-d+1}} t^d\, e^{-\beta_{k,l} t}\, \mathbb{1}_{\alpha_{i,j}\neq\beta_{k,l}} \quad (9)$$

3 Maximum

Consider a random variable being maximum of n independent random variables $X_1, \ldots, X_n$. We will denote it by $MAX(X_1, \ldots, X_n)$, its density by $max_{X_1,\ldots,X_n}$, and cdf by $MAX_{X_1,\ldots,X_n}$.

Proposition 4. *For the maximum of independent random variables $X_1, \ldots, X_n$ we have:*

$$MAX(X_1, \ldots, X_n) = MAX(X_{\delta(1)}, \ldots, X_{\delta(n)}) \quad \delta \in S_n \tag{10}$$

$$MAX(X_1, \ldots, X_n)(t) = MAX(MAX(X_1, \ldots, X_{n-1}), X_n) \tag{11}$$

Proof. These equalities follows directly from associativity and commutativity of the maximum operation.

So as in the case of the sum it is sufficient to find the density of the maximum of two random variables. Let's independent random variables X and Y have densities $f(x)$ and $g(x)$ respectively, where $f(x)$ and $g(x)$ are exponential polynomials given by:

$$f(x) = \sum_{i=0}^{\mathrm{mX}} \sum_{j=1}^{\mathrm{pX}_i} a_{i,j} x^i e^{-\alpha_{i,j} x} \mathbb{1}_{[0,\infty)}(x)$$

$$g(x) = \sum_{i=0}^{\mathrm{mY}} \sum_{j=1}^{\mathrm{pY}_i} b_{i,j} x^i e^{-\beta_{i,j} x} \mathbb{1}_{[0,\infty)}(x)$$

Proposition 5. *The cumulative distribution function of the maximum of two independent random variables X and Y with cumulative distribution functions F and G is given by:*

$$MAX_{X,Y}(t) = F(t)G(t) \tag{12}$$

Proof.

$$MAX_{X,Y}(t) = \mathbb{P}(MAX(X,Y) \leq t) = \mathbb{P}(max(X,Y) \leq t) = \mathbb{P}(X \leq t \wedge Y \leq t) = \\ = \mathbb{P}(X \leq t)\mathbb{P}(Y \leq t) = F(t)G(t)$$

So the cdf of $MAX_{X,Y}$ is given by:

$$max_{X,Y}(t) = MAX'_{X,Y}(t) = (F(t)G(t))' = F'(t)G(t)+F(t)G'(t) = f(t)G(t)+F(t)g(t)$$

Observe that for $t < 0$ we have $max_{XY}(t) = 0$, so we will continue our calculations with the assumption that $t \geq 0$.

$$max_{X,Y}(t) = \left(\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i} a_{i,j} t^i e^{-\alpha_{i,j}t}\right)\left(\sum_{i=0}^{\mathrm{mY}}\sum_{j=1}^{\mathrm{pY}_i} \frac{b_{i,j}}{\beta_{i,j}^{i+1}}\, i! - \sum_{i=0}^{\mathrm{mY}}\sum_{j=1}^{\mathrm{pY}_i}\sum_{k=0}^{i} \frac{b_{i,j}}{\beta_{i,j}^{i-k+1}}\frac{i!}{k!}\, t^k\, e^{-\beta_{i,j}t}\right) + \\ \left(\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, i! - \sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i}\sum_{k=0}^{i} \frac{a_{i,j}}{\alpha_{i,j}^{i-k+1}}\frac{i!}{k!}\, t^k\, e^{-\alpha_{i,j}t}\right)\left(\sum_{i=0}^{\mathrm{mY}}\sum_{j=1}^{\mathrm{pY}_i} b_{i,j} t^i e^{-\beta_{i,j}t}\right) = \\ = \left(\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i} a_{i,j} t^i e^{-\alpha_{i,j}t}\right)\left(\sum_{m=0}^{\mathrm{mY}}\sum_{n=1}^{\mathrm{pY}_m} \frac{b_{m,n}}{\beta_{m,n}^{m+1}}\, m! - \sum_{m=0}^{\mathrm{mY}}\sum_{n=1}^{\mathrm{pY}_m}\sum_{l=0}^{m} \frac{b_{m,n}}{\beta_{m,n}^{m-l+1}}\frac{m!}{l!}\, t^l\, e^{-\beta_{m,n}t}\right) + \\ \left(\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, i! - \sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i}\sum_{k=0}^{i} \frac{a_{i,j}}{\alpha_{i,j}^{i-k+1}}\frac{i!}{k!}\, t^k\, e^{-\alpha_{i,j}t}\right)\left(\sum_{m=0}^{\mathrm{mY}}\sum_{n=1}^{\mathrm{pY}_m} b_{m,n} t^m e^{-\beta_{m,n}t}\right)$$

After simplifying we obtain the following formula for $max_{X,Y}$ $(t \geq 0)$:

$$max_{X,Y}(t) = \left(\sum_{m=0}^{\mathrm{mY}}\sum_{n=1}^{\mathrm{pY}_m} \frac{b_{m,n}}{\beta_{m,n}^{m+1}}\, m!\right)\left(\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i} a_{i,j}\, t^i\, e^{-\alpha_{i,j}t}\right) + \\ -\left(\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i}\sum_{m=0}^{\mathrm{mY}}\sum_{n=1}^{\mathrm{pY}_m}\sum_{l=0}^{m} \frac{a_{i,j}\, b_{m,n}\, m!}{\beta_{m,n}^{m-l+1}\, l!}\, t^{l+i}\, e^{-(\alpha_{i,j}+\beta_{m,n})t}\right) + \\ \left(\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i} \frac{a_{i,j}}{\alpha_{i,j}^{i+1}}\, i!\right)\left(\sum_{m=0}^{\mathrm{mY}}\sum_{n=1}^{\mathrm{pY}_m} b_{m,n}\, t^m\, e^{-\beta_{m,n}t}\right) + \\ -\left(\sum_{m=0}^{\mathrm{mY}}\sum_{n=1}^{\mathrm{pY}_m}\sum_{i=0}^{\mathrm{mX}}\sum_{j=1}^{\mathrm{pX}_i}\sum_{k=0}^{i} \frac{a_{i,j}\, b_{m,n}\, i!}{\alpha_{i,j}^{i-k+1}\, k!}\, t^{k+m}\, e^{-(\alpha_{i,j}+\beta_{m,n})t}\right) \quad (13)$$