

5.1 - Linear Program

We interpret G with each edge colored by read mappings as a flow network, considering the read volume assigned to every (super-) edge as a flux created by the expression of the underlying supporting transcripts T . Consequently, given an edge the contribution of the supporting transcripts to the flux observed along e can be described by a linear equation

(Equation 1)

where f_i represents a factor that expresses the fraction of the respective transcript expression t_i observed between $tail_e$ and $head_e$. In the trivial case, f_i corresponds to the proportion of the interval $[tail_e; head_e]$ in comparison to the entire length of the processed transcript. The correction factor in Eq.1 is to compensate for divergence from the expectation created by stochastic sampling intrinsic to RNA-Seq experiments.

The crux of the flux is that an RNA-Seq experiment provides a series of observations on the underlying expression level t_i along the transcript body.

Following tradition in transportation problems, we model all of these observations as a system of linear equations by inferring Equation 1 on all .

Subsequently, the linear equations spanned by a locus are resolved by the objective function

(Equation 2)

Solving the linear program (Eq.2) imposed by a locus intrinsically provides an estimate for the expression level t_i of all alternative transcripts that are annotated.