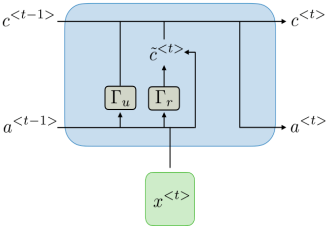
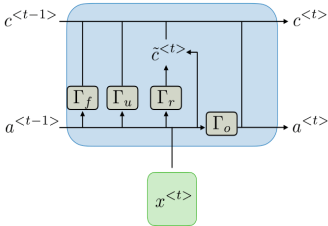


où W, U, b sont des coefficients spécifiques à la porte et σ est une sigmoïde. Les portes à retenir sont récapitulées dans le tableau ci-dessous :

Type de porte	Rôle	Utilisée dans
Porte d'actualisation Γ_u	Dans quelle mesure le passé devrait être important ?	GRU, LSTM
Porte de pertinence Γ_r	Enlever les informations précédentes ?	GRU, LSTM
Porte d'oubli Γ_f	Enlever une cellule ?	LSTM
Porte de sortie Γ_o	Combien devrait-on révéler d'une cellule ?	LSTM

GRU/LSTM – Les unités de porte récurrente (en anglais *Gated Recurrent Unit*) (GRU) et les unités de mémoire à long/court terme (en anglais *Long Short-Term Memory units*) (LSTM) appaisent le problème du gradient qui disparaît rencontré par les RNNs traditionnels, où le LSTM peut être vu comme étant une généralisation du GRU. Le tableau ci-dessous résume les équations caractéristiques de chacune de ces architectures :

	Gated Recurrent Unit (GRU)	Long Short-Term Memory (LSTM)
$\tilde{c}^{<t>}$	$\tanh(W_c[\Gamma_r \star a^{<t-1>}, x^{<t>}] + b_c)$	$\tanh(W_c[\Gamma_r \star a^{<t-1>}, x^{<t>}] + b_c)$
$c^{<t>}$	$\Gamma_u \star \tilde{c}^{<t>} + (1 - \Gamma_u) \star c^{<t-1>}$	$\Gamma_u \star \tilde{c}^{<t>} + \Gamma_f \star c^{<t-1>}$
$a^{<t>}$	$c^{<t>}$	$\Gamma_o \star c^{<t>}$
Dépendances		

Remarque : le signe \star dénote le produit de Hadamard entre deux vecteurs.

Variantes des RNNs – Le tableau ci-dessous récapitule les autres architectures RNN communément utilisées :