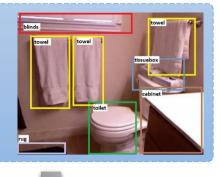
New Scene

(object
labels and
locations
detected)

Object 1



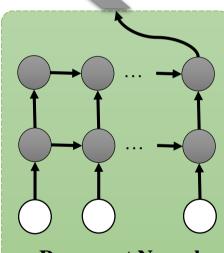
Update Model Add new context? (yes, no)

Contextualized Scene Model

(Latent Dirichlet Allocation)

Context 1 ... Context M

Feed each
context's links as
inputs at
different time
steps of RNN



Recurrent Neural

Network

(with Long Short Term

Memory units)



Hidden Units



Latent variable Visible variable

Object N

Object 2