

# Révision Examen Intra

# Introduction - Réseaux pleinement connectés

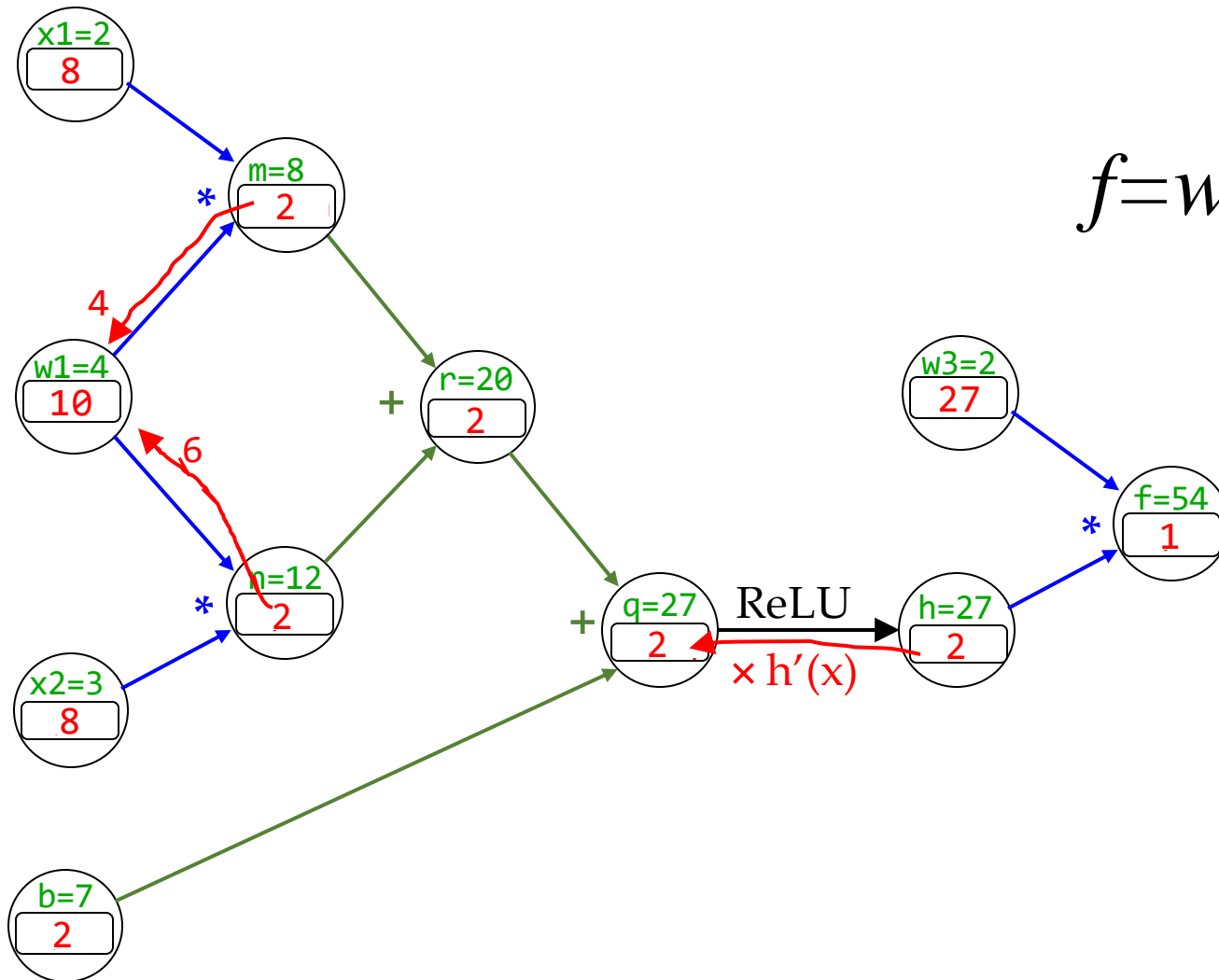
# Introduction – Fonctions d'activation

- Sigmoid
- Tanh
- Relu

# Introduction – Graphe de calcul

- Propagation avant (forward pass)
- Rétropropagation (backward pass)
- Fonctions de perte
- Dérivées en chaîne

# Introduction – Graphe de calcul



$$f = w_3 \text{ReLU}(w_1 x_1 + w_1 x_2 + b)$$

# Optimisation

- Descente de gradient stochastique
- Connaître les différents optimiseurs, leurs différences
- Taux d'apprentissage
- Stratégies d'optimisation
- Horaire d'entraînement
- Batch Norm
- Bref la section sur les méthodes pratiques...

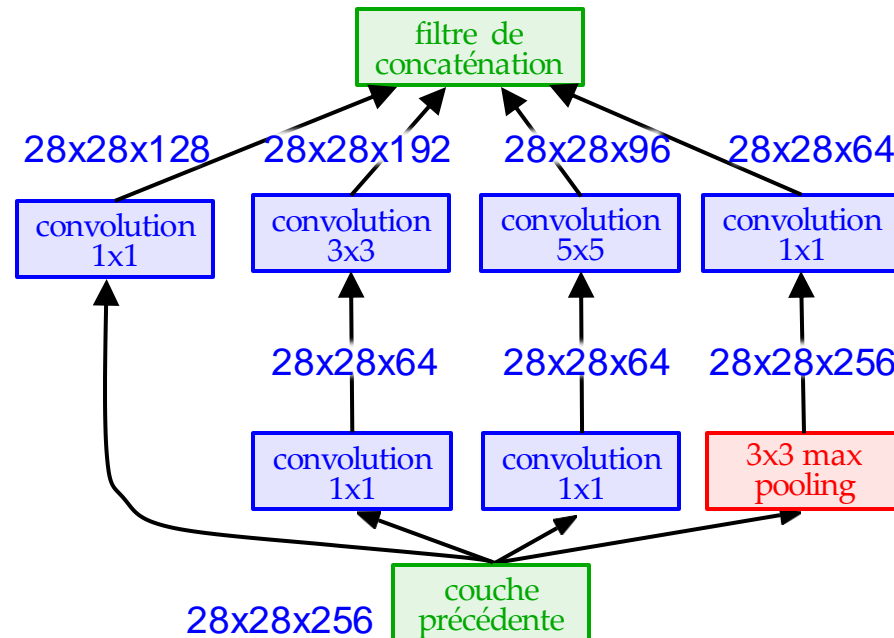
# Régularisation

- Pénalités sur les poids
- Data augmentation
- Robustesse au bruit
- Apprentissage semi-supervisé
- Apprentissage multi-tâche
- Early stopping
- Parameter tying/sharing
- Représentation sparse
- Méthodes par ensemble
- Dropout

# CNN

- Bien comprendre les filtres à convolution
- Comprendre les calculs de la convolution

## Coût en calcul



### Conv Ops:

[1x1 conv, 64]  $28 \times 28 \times 64 \times 1 \times 1 \times 256$

[1x1 conv, 64]  $28 \times 28 \times 64 \times 1 \times 1 \times 256$

[1x1 conv, 128]  $28 \times 28 \times 128 \times 1 \times 1 \times 256$

[3x3 conv, 192]  $28 \times 28 \times 192 \times 3 \times 3 \times 64$

[5x5 conv, 96]  $28 \times 28 \times 96 \times 5 \times 5 \times 64$

[1x1 conv, 64]  $28 \times 28 \times 64 \times 1 \times 1 \times 256$

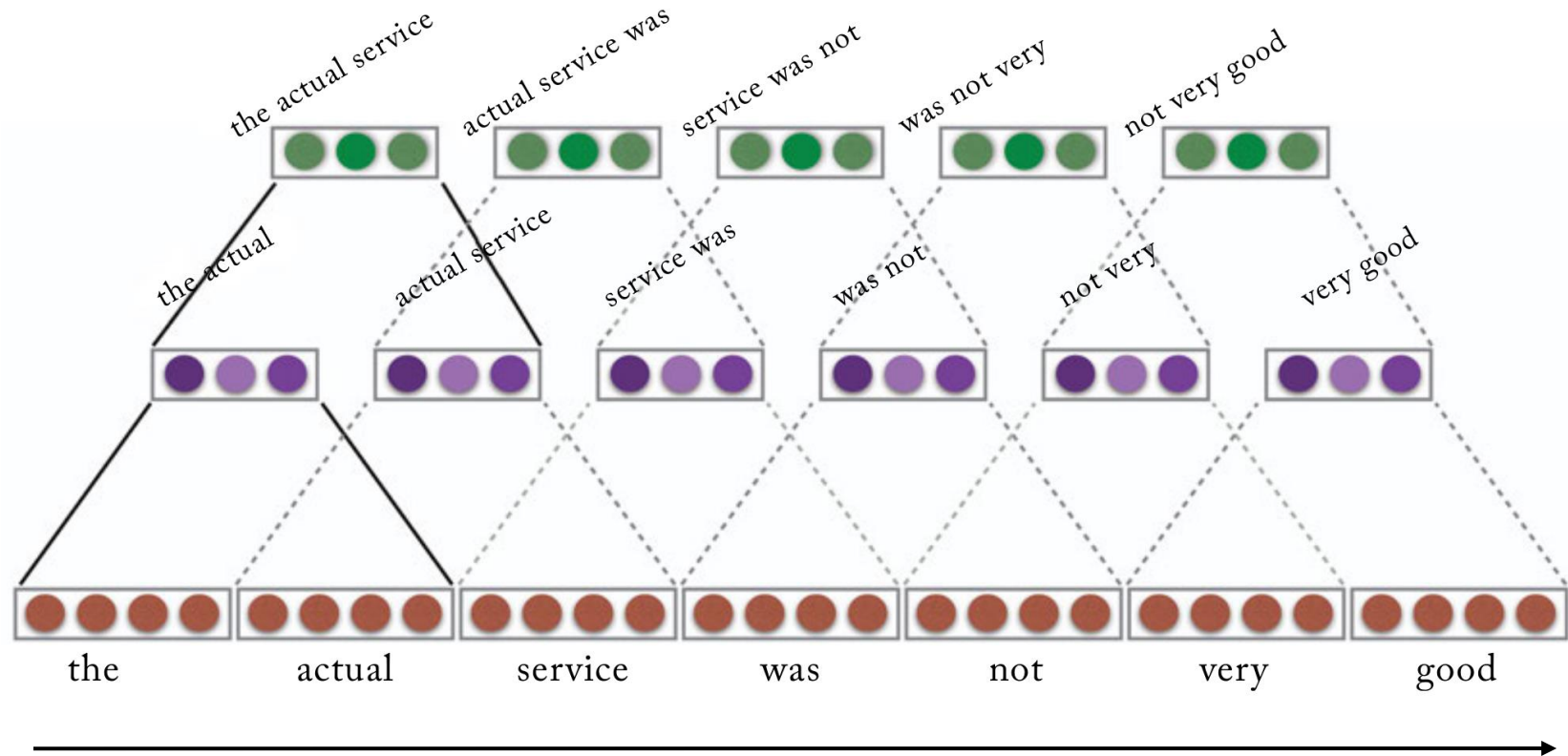
**Total: 358M ops**



# CNN

- CNN appliqués au texte

5 "mots" en sortie



7 mots en entrée

# CNN

- Différentes architectures
- Évolution au fil du temps