

УПРАВЛЕНИЕ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

глубокие нейронные сети II

А.В. Макаренко

avm@rdcn.ru

Научно-исследовательская группа «Конструктивная Кибернетика»
Москва, Россия, www.rdcn.ru

Институт проблем управления РАН
Москва, Россия

Учебный курс – Лекция 5

16 апреля 2020 г.

ИПУ РАН, Москва, Россия

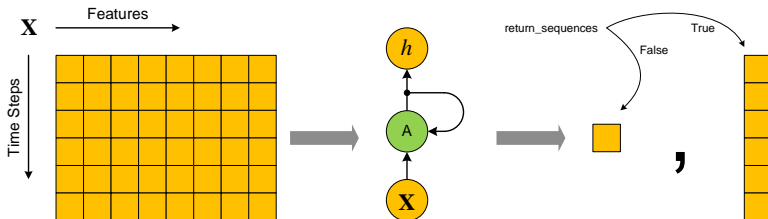
- ① Глубокие нейронные сети
- ② Заключение

Outline section

- 1 Глубокие нейронные сети
Базовая архитектура сетей II
- 2 Заключение

Рекуррентные сети I

Рекуррентная сеть (Recurrent Neural Network, RNN) в своём составе содержит нейроны охваченные обратной связью, совокупность которых формирует направленные последовательности обработки.

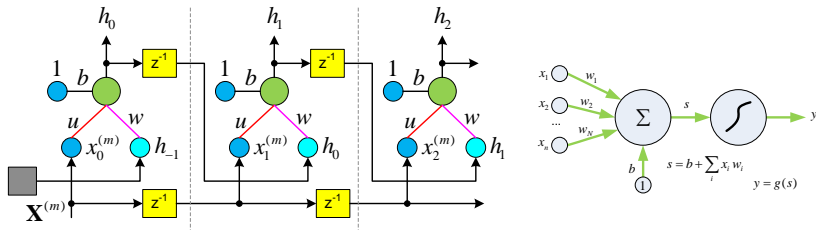


Задаётся: топология сети (topology), тип и кол-во рекуррентных нейронов в слое, параметр `return_sequences`, активационная функция.

Обучается: веса и смещения нейронов.

Рекуррентные сети II

Simple Recurrent Network I:



Сеть Элмана

$$h_k = \sigma_h(u x_k + w h_{k-1} + b_h),$$

$$y_k = \sigma_y(v h_k + b_y).$$

J.L. Elman, Finding Structure in Time.
Cognitive Science **14**:2 (1990), 179–211.

Сеть Джордана

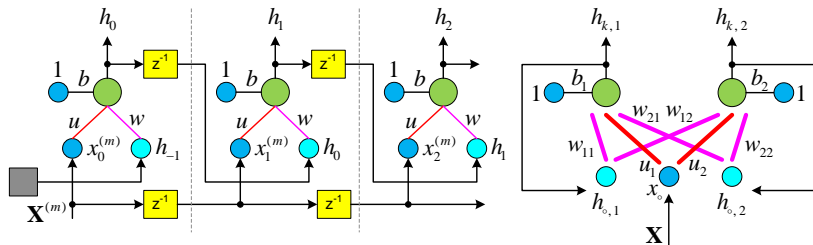
$$h_k = \sigma_h(u x_k + w y_{k-1} + b_h),$$

$$y_k = \sigma_y(v h_k + b_y).$$

M.I. Jordan, Serial Order: A Parallel
Distributed Processing Approach.
Advances in Psychology **121** (1997),
471-495.

Рекуррентные сети III

Simple Recurrent Network II:



Количество параметров в слое SimpleRNN: $MN + N^2 + N$.

M – размерность входного вектора (число признаков), $m = \overline{1, M}$.

N – число нейронов в рекуррентном слое.

Рекуррентные сети IV

$$\mathbf{h}_k = \sigma_h(\mathbf{U} \mathbf{x}_k + \mathbf{W} h_{k-1} + \mathbf{b}_h),$$

$$\mathbf{y}_k = \sigma_y(\mathbf{V} h_k + \mathbf{b}_y).$$

Рекуррентные сети IV

$$\mathbf{h}_k = \sigma_h(\mathbf{U} \mathbf{x}_k + \mathbf{W} h_{k-1} + \mathbf{b}_h),$$

$$\mathbf{y}_k = \sigma_y(\mathbf{V} h_k + \mathbf{b}_y).$$

Theorem (Full Turing of RNN)

Любые машины Тьюринга могут моделироваться полностью связанными рекуррентными сетями, созданными из нейронов с сигмоидальными функциями активации, при условии, что сеть имеет достаточное количество нейронов в скрытом слое M и достаточное число шагов временной памяти K .

H. Siegelmann, E. Sontag, Turing computability with neural nets. Appl. Math. Lett. 4(6):77-80, 1991.

Рекуррентные сети IV

$$\mathbf{h}_k = \sigma_h(\mathbf{U} \mathbf{x}_k + \mathbf{W} \mathbf{h}_{k-1} + \mathbf{b}_h),$$

$$\mathbf{y}_k = \sigma_y(\mathbf{V} \mathbf{h}_k + \mathbf{b}_y).$$

Theorem (Full Turing of RNN)

Любые машины Тьюринга могут моделироваться полностью связанными рекуррентными сетями, созданными из нейронов с сигмоидальными функциями активации, при условии, что сеть имеет достаточное количество нейронов в скрытом слое M и достаточное число шагов временной памяти K .

Полнота по Тьюрингу – характеристика исполнителя (множества вычисляющих элементов) в теории вычислимости, означающая возможность реализовать на нём любую вычислимую функцию (Алгоритм). Другими словами, для каждой вычислимой функции существует вычисляющий её элемент (например, машина Тьюринга) или программа для исполнителя, а все функции, вычисляемые множеством вычислителей, являются вычислимыми функциями (возможно, при некотором кодировании входных и выходных данных).

Рекуррентные сети IV

$$\mathbf{h}_k = \sigma_h(\mathbf{U} \mathbf{x}_k + \mathbf{W} h_{k-1} + \mathbf{b}_h),$$
$$\mathbf{y}_k = \sigma_y(\mathbf{V} h_k + \mathbf{b}_y).$$

Theorem (Universal Approximation Theorem RNN)

Любая нелинейная динамическая система может быть аппроксимирована с любой точностью рекуррентной нейронной сетью, без ограничений на компактность пространства состояний системы, при условии, что сеть имеет достаточное количество нейронов в скрытом слое (размерность n).

Рекуррентные сети IV

$$\mathbf{h}_k = \sigma_h(\mathbf{U} \mathbf{x}_k + \mathbf{W} h_{k-1} + \mathbf{b}_h),$$

$$\mathbf{y}_k = \sigma_y(\mathbf{V} h_k + \mathbf{b}_y).$$

Theorem (Universal Approximation Theorem RNN)

Любая нелинейная динамическая система может быть аппроксимирована с любой точностью рекуррентной нейронной сетью, без ограничений на компактность пространства состояний системы, при условии, что сеть имеет достаточное количество нейронов в скрытом слое (размерность n).

- Автономные ОДУ, [Funahashi & Nakamura \(1993\)](#):

$$\dot{\mathbf{s}} = \mathbf{g}(\mathbf{s}), \quad \mathbf{s}(t=0) \in S_0.$$

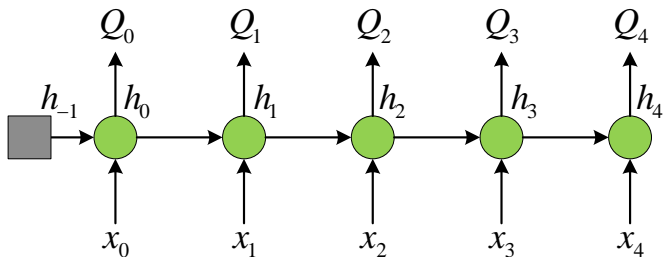
- Неавтономные ОДУ, [Chow & Li \(2000\)](#):

$$\dot{\mathbf{s}} = \mathbf{g}(\mathbf{s}) + \mathbf{f}(t), \quad \mathbf{s}(t=0) \in S_0.$$

- Примеры моделирования, ArXiv: [1512.05702](#).

Рекуррентные сети V

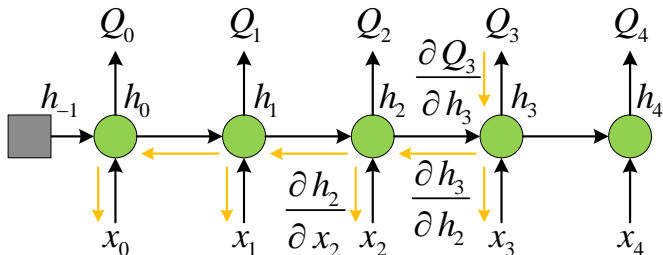
Backpropagation Through Time – алгоритм обучения рекуррентных нейросетей.



$$Q_K = \sum_{k=0}^K Q(y_k, y'_k).$$

Рекуррентные сети V

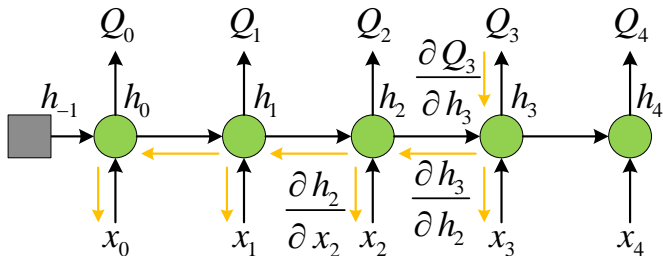
Backpropagation Through Time – алгоритм обучения рекуррентных нейросетей.



$$Q_K = \sum_{k=0}^K Q(y_k, y'_k), \quad \Delta w_{ij} = -\eta \frac{\partial Q_K}{\partial w_{ij}} = -\eta \sum_{k=0}^K \frac{\partial Q(y_k, y'_k)}{\partial w_{ij}}, \quad w_{ij} \in \{U, W, V\}.$$

Рекуррентные сети V

Backpropagation Through Time – алгоритм обучения рекуррентных нейросетей.



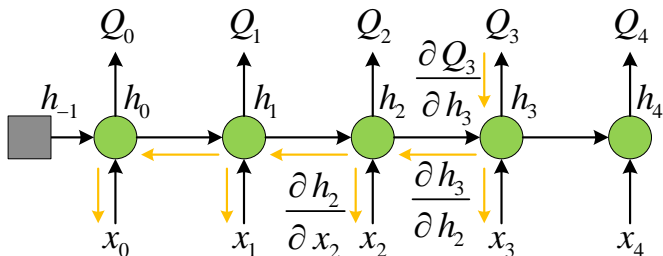
$$Q_K = \sum_{k=0}^K Q(y_k, y'_k), \quad \Delta w_{ij} = -\eta \frac{\partial Q_K}{\partial w_{ij}} = -\eta \sum_{k=0}^K \frac{\partial Q(y_k, y'_k)}{\partial w_{ij}}, \quad w_{ij} \in \{U, W, V\}.$$

Основная сложность обучения RNN (ArXiv: [1211.5063](https://arxiv.org/abs/1211.5063)):

$$\frac{\partial Q_K}{\partial \mathbf{W}} = \frac{\partial Q_K}{\partial y_K} \frac{\partial y_K}{\partial h_K} \sum_{k=0}^K \left(\prod_{j=k+1}^K \frac{\partial h_j}{\partial h_{j-1}} \right) \frac{\partial h_k}{\partial \mathbf{W}}$$

Рекуррентные сети V

Backpropagation Through Time – алгоритм обучения рекуррентных нейросетей.



$$Q_K = \sum_{k=0}^K Q(y_k, y'_k), \quad \Delta w_{ij} = -\eta \frac{\partial Q_K}{\partial w_{ij}} = -\eta \sum_{k=0}^K \frac{\partial Q(y_k, y'_k)}{\partial w_{ij}}, \quad w_{ij} \in \{U, W, V\}.$$

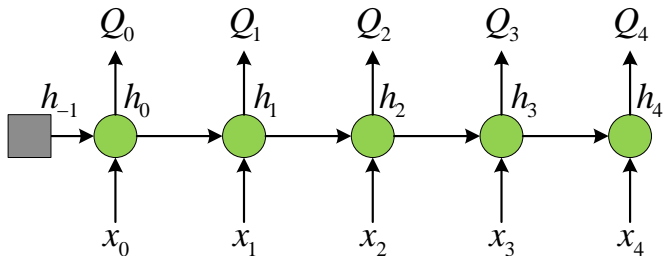
Основная сложность обучения RNN (ArXiv: [1211.5063](https://arxiv.org/abs/1211.5063)):

$$K = 128 : \quad \begin{array}{ll} w = 0.5 & G = 2.94 \cdot 10^{-39} \\ w = 1.5 & G = 3.47 \cdot 10^{22} \end{array} \quad \begin{array}{ll} w = 0.9 & G = 1.39 \cdot 10^{-6} \\ w = 1.1 & G = 1.99 \cdot 10^5 \end{array}$$

Рекуррентные сети VI

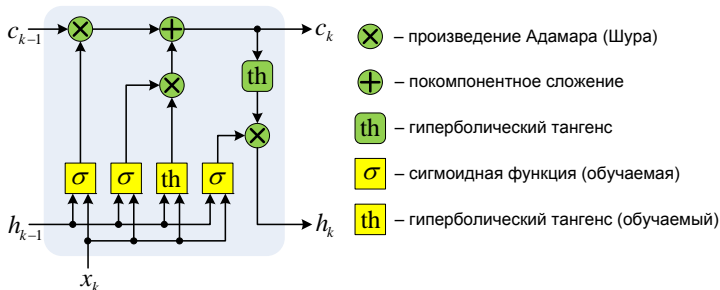
Ограничения SimpleRNN:

- Оперирование короткими контекстами.
- Сложность совмещения различных масштабов.
- Сложность обучения при больших K .



Рекуррентные сети VII

LSTM networks (сеть с долговременной и краткосрочной памятью):

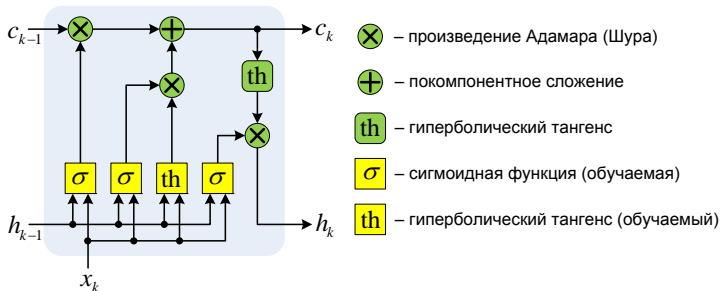


Смысл модификации: решение проблемы длинных контекстов. Запоминание информации на продолжительный срок – это одна из ключевых особенностей LSTM сетей, не требующая продолжительного обучения.

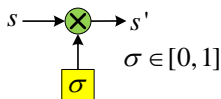
S. Hochreiter, J. Schmidhuber, Long-Short Term Memory. Neural Computation, 9(8):1735-1780, 1997. [Читать....](#)

Рекуррентные сети VII

LSTM networks (сеть с долговременной и краткосрочной памятью):

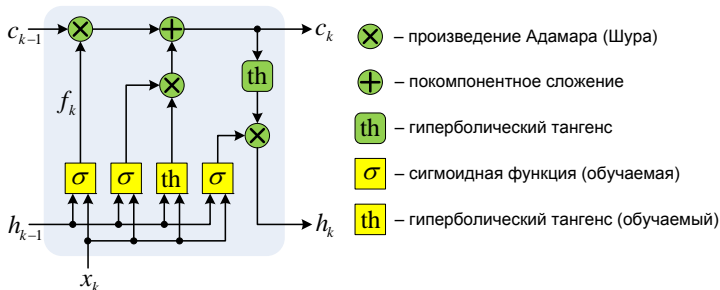


$c_{k-1} \rightarrow c_k$ – конвейер состояния ячейки – содержит только **линейные операции**. Модификация информации контролируется вентилями (gates):



Рекуррентные сети VII

LSTM networks (сеть с долговременной и краткосрочной памятью):



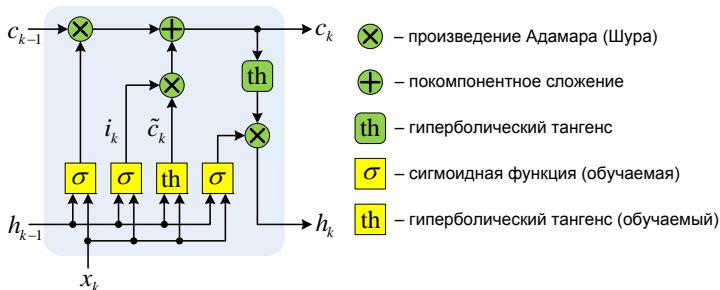
Forget gate:

$$f_k = \sigma(\mathbf{W}_f \mathbf{x}_k + \mathbf{U}_f \mathbf{h}_{k-1} + \mathbf{b}_f)$$

Интерпретация: Если тема (сцена) изменяется, то информация о старой теме (сцене) стирается.

Рекуррентные сети VII

LSTM networks (сеть с долговременной и краткосрочной памятью):



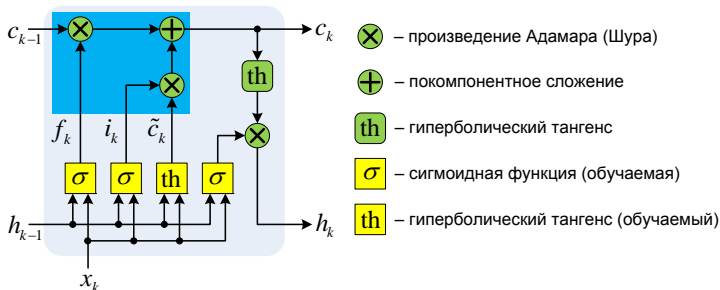
Input gate & activation:

$$i_k = \sigma(\mathbf{W}_i \mathbf{x}_k + \mathbf{U}_i \mathbf{h}_{k-1} + \mathbf{b}_i), \quad \tilde{c}_k = \tanh(\mathbf{W}_c \mathbf{x}_k + \mathbf{U}_c \mathbf{h}_{k-1} + \mathbf{b}_c)$$

Интерпретация: определяется какие значения будут обновляться и создаётся вектор кандидатов на \tilde{c}_k , которые предполагается добавить в состояние ячейки.

Рекуррентные сети VII

LSTM networks (сеть с долговременной и краткосрочной памятью):



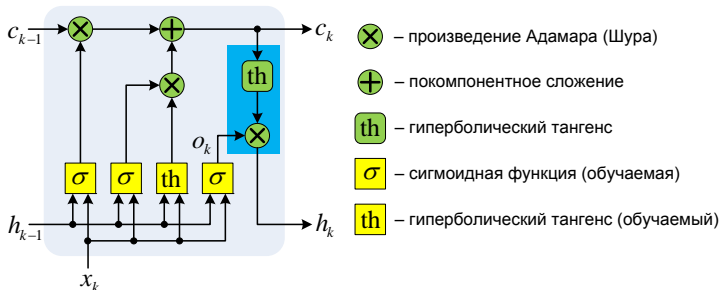
Internal state:

$$\mathbf{c}_k = \mathbf{f}_k * \mathbf{c}_{k-1} + \mathbf{i}_k * \tilde{\mathbf{c}}_k$$

Интерпретация: формируется новое состояние ячейки \mathbf{c}_k .

Рекуррентные сети VII

LSTM networks (сеть с долговременной и краткосрочной памятью):



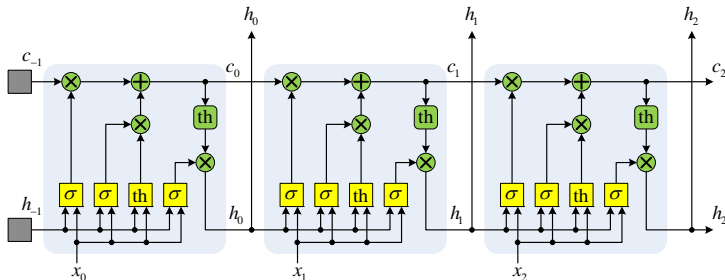
Output gate & value:

$$\mathbf{o}_k = \sigma(\mathbf{W}_o \mathbf{x}_k + \mathbf{U}_o \mathbf{h}_{k-1} + \mathbf{b}_o), \quad \mathbf{h}_k = \mathbf{o}_k * \tanh(\mathbf{c}_k)$$

Интерпретация: формируется новый выход ячейки \mathbf{h}_k .

Рекуррентные сети VIII

LSTM networks (сеть с долговременной и краткосрочной памятью):



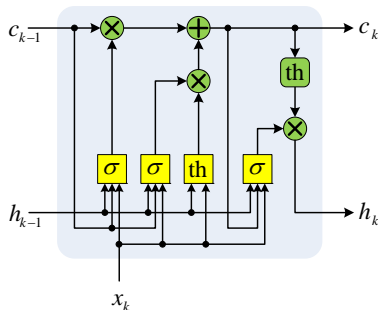
Количество параметров в слое LSTM: $4(MN + M^2 + M)$.

N – размерность входного вектора \mathbf{x} (число признаков).

M – число нейронов в рекуррентном слое, размерность: \mathbf{h} , \mathbf{c} .

Рекуррентные сети IX

LSTM networks with peephole connections:

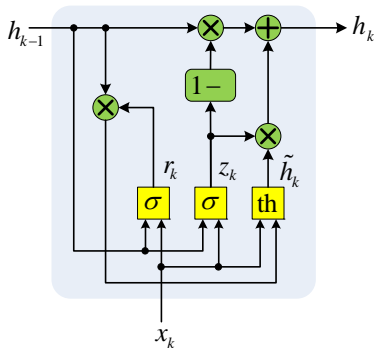


Смысл модификации: вентили наблюдают за состоянием ячейки.

F.A. Gers, J. Schmidhuber, Recurrent nets that time and count. Proceedings of the IEEE-INNS-ENNS 2000. [Читать....](#)

Рекуррентные сети X

GRU networks (рекуррентный модуль с затворами):



$$\mathbf{z}_k = \sigma(\mathbf{W}_z \mathbf{x}_k + \mathbf{U}_z \mathbf{h}_{k-1} + \mathbf{b}_z),$$

$$\mathbf{r}_k = \sigma(\mathbf{W}_r \mathbf{x}_k + \mathbf{U}_r \mathbf{h}_{k-1} + \mathbf{b}_r),$$

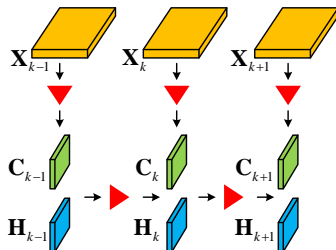
$$\tilde{\mathbf{h}}_k = \tanh(\mathbf{W}_h \mathbf{x}_k + \mathbf{U}_h [\mathbf{r}_k * \mathbf{h}_{k-1}] + \mathbf{b}_h),$$

$$\mathbf{h}_k = (1 - \mathbf{z}_k) * \mathbf{h}_{k-1} + \mathbf{z}_k * \tilde{\mathbf{h}}_k.$$

Смысл модификации: комбинируются входные вентили и вентили «забывания» в единый вентиль обновления, Кроме того объединяются скрытое и выходное состояния ячейки. ArXiv: [1406.1078](https://arxiv.org/abs/1406.1078).

Рекуррентные сети XI

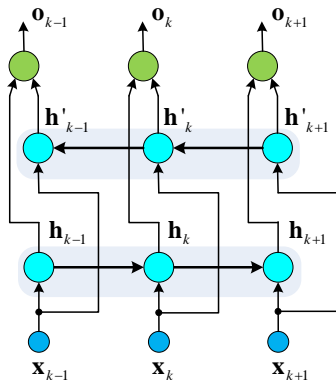
Convolutional LSTM Network (Свёрточная LSTM сеть):



Смысл модификации: «Эффективный экстрактор пространственных признаков, эффективный аппроксиматор динамики», ArXiv: [1506.04214](https://arxiv.org/abs/1506.04214).

Рекуррентные сети XII

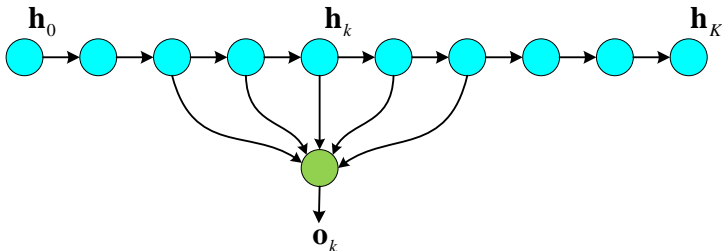
Bidirectional RNNs (Двунаправленные RNN):



Смысл модификации: «Будущее» влияет на «Прошлое», ArXiv: [1409.0473](https://arxiv.org/abs/1409.0473).

Рекуррентные сети XIII

Attention module (Модуль внимания):

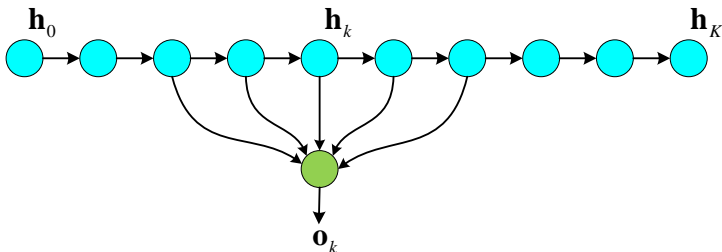


$$\mathbf{o}_k = \sum_{i=0}^{2n} a_i \mathbf{h}_{k+(i-n)}, \quad \sum_i a_i = 1.$$

Смысл модификации: Взвешенное скользящее среднее формирует для состояния ячейки контекстное окружение.

Рекуррентные сети XIII

Attention module (Модуль внимания):



$$\mathbf{o}_k = \sum_{i=0}^{2n} a_i \mathbf{h}_{k+(i-n)}, \quad \sum_i a_i = 1.$$

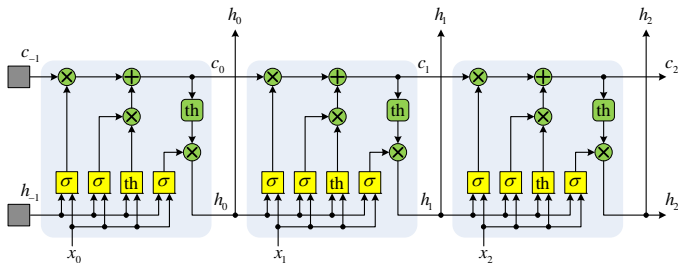
Смысл модификации: Взвешенное скользящее среднее формирует для состояния ячейки контекстное окружение.

Есть прямая связь с нейросетевыми машинами Тьюринга (Neural Turing Machine).

Рекуррентные сети XIV

Особенности обучения LSTM:

- При обратном распространении ошибки, ошибка оказывается «заперта» в памяти блока. Это называют «каруселью ошибок», которая непрерывно «скармливает» ошибку обратно каждому из вентилей, пока они не будут натренированы отбрасывать неверное значение. Таким образом, регулярное обратное распространение ошибки эффективно для тренировки в LSTM-блоках длинных контекстных последовательностей.
- Для обучения применяется алгоритм Truncated BPTT.
- Aidan Gomez, Backpropogating an LSTM: A Numerical Example. [Читать....](#)



Рекуррентные сети XV

Дальнейшие эксперименты:

- Свёрточные LSTM сети (Convolutional LSTM Network), ArXiv: [1506.04214](#).
- Решётчатая LSTM сеть (Grid LSTM), ArXiv: [1507.01526](#).
- Глубокие вентильные рекуррентные нейронные сети (Depth Gated RNNs), ArXiv: [1508.03790](#).
- Массивно-параллельная LSTM сеть (Array-LSTM), ArXiv: [1607.03085](#).
- Последовательное объединение групп LSTM слоёв (Densely Connected Bidirectional LSTM), ArXiv: [1802.00889](#).
- ...

Рекуррентные сети XV

Дальнейшие эксперименты:

- Свёрточные LSTM сети (Convolutional LSTM Network), ArXiv: [1506.04214](#).
- Решётчатая LSTM сеть (Grid LSTM), ArXiv: [1507.01526](#).
- Глубокие вентильные рекуррентные нейронные сети (Depth Gated RNNs), ArXiv: [1508.03790](#).
- Массивно-параллельная LSTM сеть (Array-LSTM), ArXiv: [1607.03085](#).
- Последовательное объединение групп LSTM слоёв (Densely Connected Bidirectional LSTM), ArXiv: [1802.00889](#).
- ...

Макаренко А.В. Глубокие нейронные сети: зарождение, становление, современное состояние // Проблемы управления, № 2 (2020), 3–19.

Outline section

① Глубокие нейронные сети

② Заключение

Контрольная работа

Задание для слушателей:

- 1 Изучить структуру рекурсивных нейронных сетей. Выписать математическое описание рекурсивного слоя. Начертить схему рекурсивного слоя.
- 2 Изучить структуру тензорных рекурсивных нейронных сетей. Выписать математическое описание тензорного слоя. Начертить схему тензорного слоя.
- 3 Дать сравнительный анализ между рекуррентными, рекурсивными и тензорными сетями (возможности, ограничения, плюсы, минусы, сложности).
- 4 Изучить статью: Shannon Egan et al. Long Short-Term Memory (LSTM) networks with jet constituents for boosted top tagging at the LHC. ArXiv: [1711.09059](#). Составить реферат. Описать плюсы и минусы решения (в контексте машинного обучения).
- 5 Изучить раздел в документации Keras: [Recurrent Layers](#).
- 6 Изучить пример [IMDB_LSTM](#). Изменить архитектуру сети, попробовать добиться максимальных показателей качества нейросети (по F1-мере).