

INVESTIGATION OF COMPUTATIONAL PROPERTIES OF
RECURRENT NEURAL NETWORKS THROUGH AN
EQUIVALENT FEED-FORWARD REPRESENTATION VIA
FEYNMAN DIAGRAMS

LORENZO TIBERI

Il lavoro di ricerca per questa tesi è stato svolto presso l'Institute for Neuroscience and Medicine n. 6 (Research Center Jülich, Germania) e supervisionato dal prof. Moritz Helias. Il progetto è stato finanziato da una borsa di studio, rilasciata da Università degli Studi di Milano, volta a supportare lavori di tesi all'estero.

MOTIVAZIONE Le reti neurali artificiali si stanno rivelando estremamente efficaci nel risolvere una grande varietà di problemi computazionali. Questi algoritmi ispirati al cervello sono costituiti da una rete di cellule computazionali (neuroni) interconnesse. In un processo chiamato apprendimento, tali reti imparano autonomamente a risolvere un dato problema modificando la propria connettività. L'abilità delle reti artificiali di organizzarsi in una macchina intelligente risulta di grande interesse per la neuroscienza. Tuttavia, questi algoritmi sono ancora molto lontani dal fornire una descrizione biologicamente plausibile dei meccanismi computazionali e di apprendimento del cervello. Le caratteristiche riscontrate in quest'ultimo che sono solitamente assenti in una rete artificiale sono: Una connettività ricorsiva, piuttosto che *feed-forward*; Una dinamica in tempo-continuo; Rumore (aleatorietà) [2]. L'obiettivo a lungo termine di questo progetto di tesi è quello di comprendere meglio come queste caratteristiche contribuiscano al potere computazionale delle reti biologiche (ma anche artificiali) e al loro meccanismo di apprendimento.

MODELLO Ho studiato la rete neurale ricorsiva descritta dalla seguente equazione differenziale stocastica

$$dx_i = -\omega_0 x_i(t) dt + \sum_{j=1}^N M_{ij} f(x_j(t)) dt - \tilde{j}_i(t) dt + D dw_i, \quad (1)$$

dove $x_i(t)$ è lo stato del neurone i -esimo al tempo t . L'elemento di matrice M_{ij} modella la sinapsi che trasporta il segnale mandato dal neurone j -esimo al neurone i -esimo. Il segnale è una funzione f dello stato neuronale. I valori di M possono cambiare durante il processo di apprendimento, modellando dunque la plasticità sinaptica. La frequenza caratteristica ω_0 descrive un termine di perdita; \tilde{j} è uno stimolo esterno ricevuto dal cervello; dw è un incremento di Wiener che descrive rumore bianco all'interno di ogni neurone. L'intensità del rumore è parametrizzata da $D > 0$.

METODO Il formalismo di Martin-Siggia-Rose-DeDominicis-Janssen [4, cap. 7] può essere utilizzato per descrivere in modo equivalente l'equazione (1) in termini di un funzionale generatore dei momenti, mediante il quale è possibile calcolare la statistica dello stato neuronale $x(t)$. Questa riformulazione mi ha permesso di studiare il sistema usando i metodi ben consolidati della teoria statistica dei campi.

RISULTATI L'utilizzo dell'approccio appena descritto in campo *machine learning* rappresenta un'idea completamente nuova e inesplorata. Dunque ho

prima di tutto definito i dettagli tecnico-matematici del metodo e derivato dei risultati formali che mi hanno permesso di individuare tre possibili direzioni di indagine.

Direzione 1 In assenza di rumore, un algoritmo fornito da Pearlmutter [6] permette di addestrare la mia rete. Un problema ben noto di questo algoritmo è quello dei *vanishing gradients* [1], che impedisce alla rete di riconoscere caratteristiche su lunga scala temporale del segnale in ingresso. Ho riformulato le equazioni di Pearlmutter in un linguaggio di teoria dei campi, rivelandone in particolare il legame con la funzione di risposta lineare del sistema. Questa funzione è direttamente legata alle proprietà mnemoniche della rete, dal momento che quantifica quanto, col passare del tempo, il sistema abbia memoria di una piccola perturbazione avvenuta nel passato. Dunque uno studio dettagliato della sua struttura potrebbe aiutarci ad affrontare il problema menzionato.

direzione 2 Ci sono molte ragioni per pensare che il rumore presente nelle reti biologiche implementi una regolarizzazione intrinseca, potenziando dunque la capacità del sistema di generalizzare quanto imparato in fase di apprendimento [5]. Ho esteso le equazioni di Pearlmutter al caso rumoroso ($D > 0$) avvalendomi del formalismo dell'azione efficace (anche chiamata funzione di vertice) [4, chap. 11]. La tecnica della *loopwise expansion* permette di espandere l'azione efficace intorno al caso privo di rumore, includendo sistematicamente le fluttuazioni statistiche come correzioni dell'ordine di potenze di D . Questo potrebbe aiutarci a studiare l'effetto del rumore in fase di apprendimento.

direzione 3 La ragione del successo e della popolarità delle reti artificiali *feed-forward* risiede anche nella loro semplicità matematica. Ho ottenuto una rappresentazione equivalente della mia rete fortemente ricorsiva in termini di una *feed-forward*, semplificando dunque l'analisi matematica della mappatura input-output realizzata dal sistema. Questa rappresentazione è stata ottenuta considerando uno sviluppo perturbativo in diagrammi di Feynman per piccole non-linearità in f (Figura 1.a).

Ho infine deciso di ottenere risultati meno formali e più pratici in quest'ultima direzione. Ho considerato una $f(x) = x + \frac{\alpha}{2}x^2$ leggermente non lineare, dove $\alpha \ll 1$, e troncato l'espansione all'ordine $\mathcal{O}(\alpha)$. Questa è la prima cosa che si può fare per indagare l'effetto di una f non lineare sul potere computazionale del sistema. Ho mostrato che l'operazione che la rete esegue su di un input statico $\vec{j}(t) = \vec{j}\delta(t)$ è la stessa di un *quadratic threshold gate*. Ho quindi calcolato la sua *storage capacity*, generalizzando un risultato di Gardner *et al.* per il *linear threshold gate* [3]. Questa quantità è un indicatore del potere computazionale del sistema. Ho trovato che la capacità cresce quadraticamente con il numero di neuroni, invece che linearmente come nel caso $\alpha = 0$ (Figura 1.b). Dunque, anche per piccole non-linearità di f , il guadagno in potere computazionale ottenuto dal sistema diventa comunque molto grande per un grande numero di neuroni. In neuroscienza teorica, i modelli studiati analiticamente sono solitamente lineari. Il mio risultato mostra che, quando vogliamo pensare al cervello come ad un sistema computazionale, anche piccole non-linearità diventano rilevanti e non possono essere trascurate per un grande numero di neuroni. I risultati ottenuti analiticamente sono stati da me confermati anche con metodi numerici.

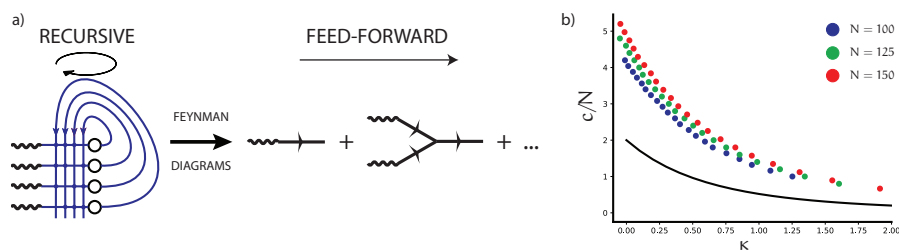


Figura 1: **a)** Rappresentazione mediante diagrammi di Feynman di una rete ricorsiva (sinistra) in termini di una *feed-forward* (destra). L'input \vec{j} è rappresentato da linee ondulate; Le linee orientate rappresentano propagatori che trasportano l'input verso l'output. Lo sviluppo diagrammatico corrisponde ad una $f(x) = x + \frac{\alpha}{2}x^2$ ed è troncato a $\mathcal{O}(\alpha)$. Si noti che in una rete ricorsiva l'input percorre dei loop all'interno del sistema, mentre con i diagrammi di Feynman viene processato in una casacata di azioni sequenziali prive di loop. **b)** Capacità C (divisa per il numero di neuroni N) in funzione del margine di classificazione κ (si veda [9] per una definizione). I punti rappresentano risultati numerici per i casi $\alpha = 0.1$ e $N = 100, 125, 150$. La linea continua rappresenta il caso lineare $\alpha = 0$, per qualsiasi N . Si noti come, nel caso non lineare, la capacità C cresce più che linearmente al crescere di N .

BIBLIOGRAPHY

- [1] Y. Bengio, P. Simard, and P. Frasconi. "Learning long-term dependencies with gradient descent is difficult." In: *IEEE Transactions on Neural Networks* 5.2 (Mar. 1994), pp. 157–166. DOI: [10.1109/72.279181](https://doi.org/10.1109/72.279181).
- [2] Valentin Braitenberg and Almut Schüz. *Anatomy of the Cortex: Statistics and Geometry*. Berlin, Heidelberg, New York: Springer-Verlag, 1991.
- [3] E Gardner. "Maximum Storage Capacity in Neural Networks." en. In: *Europhysics Letters (EPL)* 4.4 (Aug. 1987), pp. 481–485. DOI: [10.1209/0295-5075/4/4/016](https://doi.org/10.1209/0295-5075/4/4/016). (Visited on 04/08/2019).
- [4] Moritz Helias and David Dahmen. "Statistical field theory for neural networks." In: *arXiv:1901.10416 [cond-mat]* (Jan. 2019). arXiv: 1901.10416. URL: <http://arxiv.org/abs/1901.10416> (visited on 05/21/2019).
- [5] Mantas Lukoševičius and Herbert Jaeger. "Reservoir computing approaches to recurrent neural network training." In: *Computer Science Review* 3.3 (Aug. 2009), pp. 127–149. DOI: [10.1016/j.cosrev.2009.03.005](https://doi.org/10.1016/j.cosrev.2009.03.005). URL: <https://doi.org/10.1016/j.cosrev.2009.03.005>.
- [6] Barak A. Pearlmutter. "Learning State Space Trajectories in Recurrent Neural Networks." In: *Neural Computation* 1.2 (June 1989), pp. 263–269. DOI: [10.1162/neco.1989.1.2.263](https://doi.org/10.1162/neco.1989.1.2.263). URL: <https://doi.org/10.1162/neco.1989.1.2.263> (visited on 06/20/2019).