

La traduzione nell'era dei *Big Data*

È solo la lingua che fa eguali. Eguale è chi sa esprimersi e intende l'espressione altrui. Che sia ricco o povero importa meno. Basta che parli

Don Lorenzo Milani

Caterina Milone, Aurora Paladini, Lorenzo Pellegrino

Scuola Superiore ISUF

Nella società dell'informazione e nell'era del digitale può essere particolarmente interessante approfondire la realtà dei *Big Data* come strumento e come risorsa essenziale per il progresso e l'evoluzione culturale e scientifica. Dopo aver dato una definizione di *Big Data* e aver menzionato alcuni esempi di ciò, in particolare quelli legati alle traduzioni di testi da lingue diverse, si parlerà di processamento delle lingue nel cervello umano. Si mostrerà come i recenti sviluppi tecnologici abbiano inciso significativamente nel campo della traduzione, portando la figura professionale del traduttore a modificare e integrare le proprie abilità e competenze in cambio di un sistema in grado di garantire una maggiore produttività ed efficienza. In tale direzione si approfondirà il sistema di traduzione automatica neurale sviluppato da Google, confrontandone poi, alcuni risultati con le medesime traduzioni effettuate da un traduttore umano

Introduzione

La conoscenza necessaria per comprendere i continui cambiamenti del mondo è in gran parte costituita da informazioni espresse tramite mezzi linguistici. In particolare, la padronanza delle lingue e la traduzione da una all'altra sono essenziali per l'accesso e la condivisione di conoscenze in diverse parti del globo.

La padronanza del linguaggio non è oggi un'esclusiva dell'*homo sapiens sapiens*, poiché anche le macchine riescono a processare moli significative di informazioni per apprendere, comprendere e produrre linguaggio. Cosa avviene nel cervello quando elaboriamo una frase e la traduciamo? In un mondo in cui la tecnologia permette di elaborare il linguaggio e tradurlo, qual è il ruolo di un traduttore umano? Quali algoritmi può utilizzare oggi una macchina per tradurre una frase? I prossimi paragrafi hanno l'obiettivo di offrire degli spunti per poter rispondere a queste domande.

I Big Data

L'apprendimento del linguaggio e la sua elaborazione da parte delle macchine richiedono l'utilizzo dei *Big Data*. Sebbene questa espressione sia sempre più diffusa, non c'è una definizione univoca di cosa si intenda con essa. In questa sede si indicherà con *Big Data* il **capitale di informazioni caratterizzato da un volume, una velocità ed una varietà tali da richiedere tecnologie e metodi analitici specifici per la sua trasformazione in valore**. La definizione scelta individua quattro principali ambiti legati ai *Big Data*: le informazioni, la tecnologia, i metodi di analisi e l'impatto prodotto da essi nella società.

Per quanto riguarda il primo aspetto, l'affermazione dei *Big Data* è legata alla produzione, alla condivisione e all'uso massivo di dati digitali. La digitalizzazione, cioè la trasformazione dei segnali da analogici a digitali, è diventata una pratica diffusa negli anni '90; essa ha permesso l'organizzazione delle informazioni in set più strutturati, offrendo un diverso punto di vista per studiare trend e pattern rilevanti. Inoltre, la diffusione di dispositivi personali ha favorito l'aumento della disponibilità di informazioni; questi dispositivi, infatti, sono dotati di sensori che permettono la digitalizzazione e di una connessione ad Internet che permette ai dati di essere raccolti, trasformati ed organizzati. Per descrivere le caratteristiche dei *Big Data* vengono solitamente utilizzate "le 3 V" di Laney [1]: Volume, Velocità e Varietà, riferendosi all'aumento di queste dimensioni. Ad esse sono state aggiunte successivamente la Veridicità, il Valore, la Complessità e l'*unstructuredness*, caratteristica dei dati non strutturati come i video.

Il secondo aspetto, conseguente dalla complessità dei dati, riguarda i sistemi di processamento: essi devono avere capacità computazionali, di immagazzinamento e di trasmissione superiori rispetto ai sistemi utilizzati mediamente nell'ambito delle tecnologie dell'informazione. La comunicazione deve raggiungere velocità maggiori (tendendo idealmente al tempo reale) per quantità maggiori di informazioni, secondo Kitchin [2] dell'ordine dei terabytes o dei petabytes.

Il terzo aspetto riguarda i metodi adoperati per analizzare i dati. La sfida dell'analisi, infatti, è stata superata solo recentemente grazie

a tecniche provenienti dalla ricerca sull'intelligenza artificiale e a sistemi in grado di costruire modelli predittivi. Tra le procedure più utilizzate, Manyika e Chen segnalano: *Cluster analysis, Genetic algorithms, Natural Language Processing, Machine learning, Neural Networks, Predictive modelling, Regression Models, Social Network Analysis, Sentiment Analysis, Signal Processing* e *Data Visualization*.

Infine, viene considerato l'impatto che la gestione e l'utilizzo dei *Big Data* hanno su molti ambiti della società. Boyd e Crawford [1] hanno definito i *Big Data* come un fenomeno culturale oltre che tecnologico ed accademico legato alla credenza che grandi set di dati offrano una maggiore oggettività. Mayer-Schönberger [1] e Cukier [1] affermano che l'impiego di questi dati abbia prodotto cambiamenti nel modo di analizzare le informazioni per migliorare la comprensione e l'organizzazione della società. L'adattabilità dei *Big Data* alle esigenze di diversi domini scientifici ed industriali ne ha permesso l'applicazione per la risoluzione di problemi di vario genere. Tra le varie applicazioni si annoverano medicina personalizzata, simulazioni scientifiche, analisi di dati provenienti dai social network e processamento del linguaggio. In particolare per quanto riguarda il linguaggio alcune funzioni molto utilizzate sono il riconoscimento e la traduzione di testi, quest'ultima costituirà il tema centrale del presente articolo.

Il linguaggio umano, le lingue e la macchina "cervello"

La traduzione, un'abilità concomitante al bilinguismo, consiste nella riformulazione di un input formulato in una lingua sorgente (SL) verso una lingua target (TL). Prima di spiegare come avviene la traduzione di una frase grazie alle nuove tecnologie è opportuno illustrare brevemente ciò che permette ad un essere umano di compiere questa operazione. I requisiti per compiere una traduzione si potrebbero riassumere in: conoscenza della SL e della TL, capacità di leggere e comprendere la frase iniziale e operare una trasposizione del contenuto nella seconda lingua.

Innanzitutto occorre precisare che ciò a cui in inglese ci si riferisce col termine *language cor-*

risponde in italiano a due termini distinti: linguaggio e lingua. Il linguaggio viene definito, in psicolinguistica, come una "funzione cognitiva che consente di acquisire e usare una lingua". Quest'ultima è un "sistema di simboli e regole per combinarli così da poter generare un numero infinito di possibili messaggi e significati". Solitamente i linguisti parlano di linguaggio al singolare, in quanto capacità caratteristica degli esseri umani, e di lingue anche al plurale, poiché esse indicano forme specifiche di comunicazione diverse nelle varie comunità.

Una caratteristica delle lingue è avere una struttura gerarchica in cui si possono individuare più livelli. Il livello più "elementare" è quello dei suoni (fonemi), seguito da quello dei morfemi e delle parole, quello delle frasi (sintassi) ed infine quello dei significati (semantica). Per ognuno di questi livelli gli individui acquisiscono delle conoscenze (chiamate competenze, in linguistica) che permettono loro di parlare e comprendere una determinata lingua e fanno parte della "grammatica" immagazzinata nella loro mente.

Per acquisire la lingua madre (L1), i neonati non partono da regole precise, ma dai *dati linguistici primari*, cioè dagli "enunciati che il mondo circostante pronuncia" attorno a loro. Inizialmente sono in grado di discriminare tutti i tipi di suoni, col passare dei mesi smettono di percepire differenze non importanti nella lingua che stanno acquisendo e imparano che alcune combinazioni di suoni sono molto più probabili di altre nonostante il "problema della povertà dello stimolo", per cui il modello di lingua a cui sono esposti i neonati è solitamente povero, disturbato, variabile e discontinuo, essi riescono ad acquisire rapidamente una lingua uniforme e solida. Oltre a distinguere i singoli fonemi i bambini imparano a riconoscere le parole (*segmentazione*), le memorizzano e le associano ad un significato. Essi iniziano a produrre lallazioni, cioè a ripetere componenti sillabiche senza significato, solitamente verso i 6 mesi; dopo i 18 mesi comprendono circa 150 parole e ne usano 50; dopo 1 o 2 anni hanno acquisito i suoni il ritmo e l'accento della lingua a cui sono stati esposti. Per l'apprendimento della lettura sono state individuate tre fasi: inizialmente i bambini riconoscono alcune parole familiari nella loro totalità come delle immagini, successivamente appren-

dono le regole di conversione fonema-grafema riconoscendo le parole una lettera/sillaba per volta, infine le tracce mnestiche delle parole scritte vengono associate ai significati astratti e la lettura diventa automatica.

Sebbene l'acquisizione della L1 sia semplice ed intuitiva per un bambino, lo studio di come avvenga questo processo non è altrettanto facile; esso è, infatti, complesso e multi-modale, coinvolge capacità computazionali, cognitive e sociali. Non si conoscono ancora esattamente i meccanismi cerebrali che permettono di imparare una lingua; gli studiosi sono, tuttavia, abbastanza concordi nell'affermare l'esistenza di periodi critici per l'apprendimento fonetico, lessicale e sintattico legati all'età e allo sviluppo del cervello. La Figura 1 mostra, con delle fasce colorate, lo sviluppo tipico delle abilità linguistiche proprio in funzione di questi due fattori: sull'asse delle ordinate viene rappresentato il volume del cervello (100 corrisponde al valore medio per un adulto), mentre su quello delle ascisse sono rappresentati gli anni.

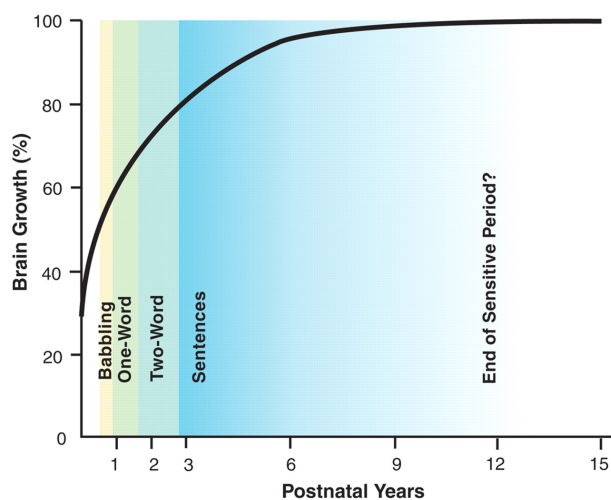


Figura 1: Crescita del cervello e tappe di acquisizione della prima lingua (L1). Fonte: [3] pag. 816

La nozione di **periodo sensibile** è legata alla perdita di flessibilità nella riorganizzazione cerebrale dopo la pubertà (circa 12 anni) ed è legata alle differenze riscontrate nell'apprendimento di una seconda lingua (L2). Quest'ultimo si mostra molto più variabile rispetto all'acquisizione della L1 e non è ancora chiaro se sia basato sugli stessi meccanismi. Una L2 può essere appresa in ogni periodo della vita, tuttavia raramente si raggiunge in essa un livello di abilità molto alto se

l'apprendimento avviene dopo i periodi sensibili (Figura 2).

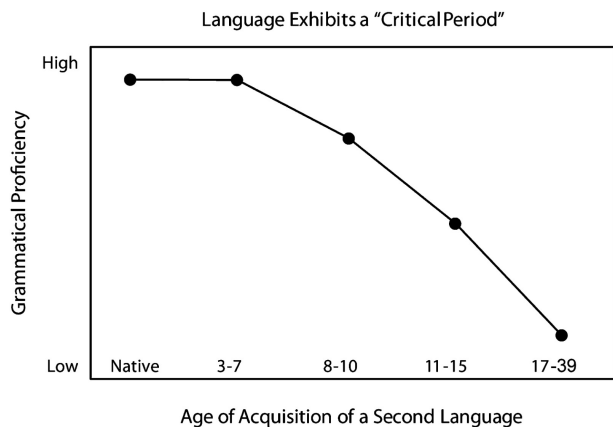


Figura 2: Relazione tra età di acquisizione di una seconda lingua (L2) e capacità linguistiche. Fonte: [4] p. 716.

Dopo aver immagazzinato informazioni relative alla L1 e ad una L2 è possibile leggere una frase e tradurla ricorrendo al contesto per risolvere eventuali ambiguità.

Tradurre consiste nell'instaurare una stretta interrelazione tra testo, lingua e mente. Non si tratta semplicemente di ricercare un equivalente di un'unità linguistica. Si tratta di trasferire significato da un sistema linguistico ad un altro, il che attiva una serie di meccanismi complicati. Alla base della traduzione si trova la rappresentazione mentale del testo nella mente del traduttore attraverso il canale della lettura. Questo passaggio attiva la fase di decodifica del prototesto (testo di partenza) da parte della macchina neurale (la psiche) che porterà alla successiva ricodifica nella forma del metatesto (testo di arrivo).

Una fase necessaria alla comprensione di una frase scritta è l'identificazione delle parole tramite la percezione visiva per poter, in seguito, accedere ai loro significati. Alcuni studi sul deficit di categorizzazione hanno permesso di ipotizzare che la rappresentazione semantica dipenda dalla combinazione di diverse informazioni percettive e funzionali. I significati delle parole, dunque, sarebbero distribuiti in più aree cerebrali tramite delle reti. Si è ipotizzato che per comprendere una parola scritta si possa accedere direttamente al significato dall'informazione visiva, oppure tramite la sua rappresentazione fonologica. Uno studio di *neuroimaging* ha permesso di rilevare l'importanza di due regioni dell'emisfero sini-

stro, di solito maggiormente coinvolto nel procesamiento del linguaggio, legate ai due processi di lettura citati. La prima regione è situata nell'area temporo-parietale ed è legata all'apprendimento che coinvolge la via fonologica; la seconda si trova nella zona occipito-temporale e connette le informazioni visive al significato (Figura 3).

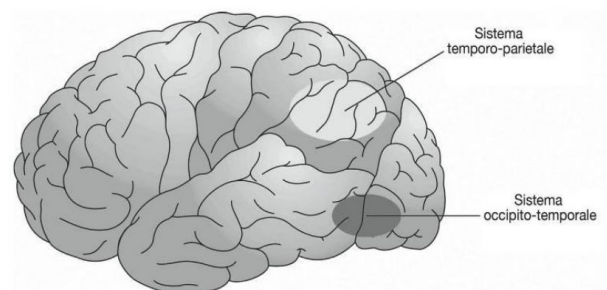


Figura 3: Regioni cerebrali importanti per la lettura. Fonte: [5] p.475.

Per quanto riguarda il cervello in azione durante la traduzione, da alcuni studi di *imaging* su soggetti bilingue sembra che i substrati neurali coinvolti nella traduzione dipendano dall'unità di testo considerata (parola, frase o testo) e dalla direzione della traduzione; l'emisfero sinistro è coinvolto esclusivamente o in maniera predominante nella traduzione di tutti i tipi di unità; l'area di Broca, situata nel lobo frontale, sembra avere un ruolo in tutti i compiti di traduzione; non sono state identificate regioni specifiche legate in modo esclusivo alla traduzione. Fattori come l'età di acquisizione, il tipo di lingue conosciute e la frequenza d'uso della traduzione nella vita quotidiana potrebbero influenzare l'organizzazione cerebrale, legata a questa funzione, nei diversi individui. Sebbene gli studi sulla localizzazione delle aree funzionali per la traduzione non offrano ancora risultati chiari, le strutture subcorticali sembrerebbero fondamentali.

La traduzione tra storia e tecnologia

Tradurre, dal latino *traducere*, it. *trasportare*, pone le sue radici lontano nella storia, a quando, per scopi commerciali, si rendeva necessaria la presenza dell'interprete in modo da poter stipulare accordi tra le parti. Nonostante ciò, la traduzione come filone di ricerca, la traduttolo-

gia, ha preso piede soltanto alla fine degli anni '50, quando vennero pubblicati i primi studi e le primissime teorie sui processi traduttivi. Sembra quasi un paradosso che proprio negli stessi anni sia stata sperimentata una delle prime applicazioni di carattere non numerico, la *Machine Translation* (MT), ovvero la traduzione automatica. I primi sistemi sviluppati in questi anni, basati sulle strutture grammaticali delle lingue in oggetto, hanno riscosso scarso successo a causa dei risultati di bassa qualità che fornivano. Una prima svolta si è osservata verso la fine degli anni '90, quando ricercatori IBM hanno creato un corpus linguistico partendo da testi paralleli in inglese e in francese prodotti dal Parlamento canadese: questo ha permesso per la prima volta il raccoglimento di dati per l'elaborazione di statistiche riguardanti i record di traduzione e di correlazione e l'introduzione di un modello probabilistico per la traduzione automatica. La vera svolta, tuttavia, risiede nella rivoluzione delle reti neurali a cui abbiamo assistito negli ultimi anni, grazie agli sforzi combinati negli studi in ambito delle neuroscienze e dell'ingegneria informatica. Ricorrendo all'impiego di modelli sequenziali basati sul *deep learning*, la ricerca sulla traduzione automatica si è concentrata su una particolare tipologia di rete neurale, detta ricorrente, dotata di unità computazionali con una memoria *short-long term* aumentata che permette di mantenere informazioni sul contesto di riferimento dall'inizio fino alla fine di una determinata frase. Quindi, il successo della linguistica computazionale (NLP) che si è osservato nel corso degli ultimi 20 anni è riconducibile ad almeno quattro fattori, quali:

- il significativo aumento della potenza computazionale;
- la vasta disponibilità di dati linguistici (*Big Data*);
- lo sviluppo di sistemi di *Machine Learning* efficienti;
- l'avanzamento della ricerca nei campi della psicolinguistica e della neurolinguistica.

Nuovi strumenti per l'industria linguistica

Big Data e nuove tecnologie hanno rivoluzionato ogni settore e quindi anche l'industria delle lingue, portando a cambiamenti non di poco conto nel mondo della traduzione. Già prima che si arrivasse a ciò, il lavoro del traduttore era sostanzialmente cambiato quando sono stati introdotti i cosiddetti *CAT Tools* (*Computer Aided-Translation Tools*) all'inizio degli anni '90 dall'azienda tedesca Trados GmbH. Sviluppati per far fronte alle esigenze legate alla globalizzazione, questi software hanno permesso ai traduttori di lavorare in modo più produttivo ed efficiente. Al centro di questi ambienti di traduzione si trovano le memorie di traduzione (*Translation Memory*), ovvero sistemi che permettono di mettere a confronto testi paralleli in cui la traduzione è prodotta dal traduttore stesso. Accanto alle memorie di traduzione, altri strumenti utili sono i *corpora* linguistici, raccolte di testi in una stessa lingua e/o riguardanti lo stesso argomento che possono essere interrogati ai fini della ricerca terminologica. Come risultato del processo di innovazione tecnologica, ricorrere ai *CAT Tools* non significa più soltanto fare ricorso alle memorie di traduzione, ma anche a una serie di strumenti come glossari, supporto alla localizzazione, controllo ortografico e traduzione automatica. A partire da questi strumenti, lo scenario a cui assistiamo oggi è molto più articolato. Grazie ai processi di digitalizzazione e alle piattaforme *Cloud* si apre una nuova era in cui il lavoro di ogni singolo traduttore può essere condiviso in rete, andando a creare un'ingente mole di dati mirata a rendere il processo traduttivo ancora più efficiente. I vantaggi dei *Big Data* non riguardano solo i singoli, ma si estendono alle agenzie di traduzione che saranno in grado di monitorare la domanda di servizi linguistici, la qualità del prodotto finale e, naturalmente, predire tendenze future. Una competenza sempre più richiesta nel mondo della traduzione tecnica è quella del *Post-editing*, strettamente connesso alla *Machine Translation*. La rete neurale di Google (GNMT), lanciata nel 2016, è tra le più avanzate con una percentuale di successo pari all'85% in riferimento alle traduzioni da e verso la lingua inglese. Al fine di aumentare ulteriormente la produttività, prende

sempre più piede il PEMT (*Post-editing Machine Translation*), almeno per quanto riguarda i testi in lingua inglese. Con il PEMT, il testo di partenza viene fornito al servizio di traduzione automatica che lo traduce in tutte le sue parti. In seguito, l'output viene affidato ad un traduttore umano per la correzione degli errori e la revisione per ottenere il testo di arrivo, il prodotto finale del processo traduttivo.

La traduzione automatica di Google

Per anni i sistemi di traduzione più all'avanguardia si sono poggiati sulla sostituzione parola-parola, senza preoccuparsi del contesto. Per questo si necessitava di avere la traduzione in più lingue di ogni parola (anche separatamente dal contesto). Sembra ragionevole pensare che questo tipo di traduzione non risulti particolarmente efficace e non restituisca il senso della frase da tradurre, dal momento che non tutte le lingue posseggono le stesse strutture grammaticali o sintattiche. Questo è il primo motivo per cui si è passati a sistemi di traduzione più avanzati, come la SMT (*Statistical Machine Translation*), che con l'avvento delle reti neurali è stata inglobata nella NMT (*Neural Machine Translation*). Diversi studi hanno proposto una rete neurale detta *RNN Encoder-Decoder* che è in grado di apprendere ad allineare i testi, catalogarli e tradurli.

Si pensi ad un bambino che apprende una lingua: gli si possono insegnare alcune parole e alcune regole grammaticali di base per costruire le frasi ma, proseguendo gli studi, sopravvengono una serie di eccezioni e casi particolari di cui si deve tenere conto. È questo il momento in cui si deve accostare allo studio del *vocabolario*, lo studio dei *testi originali*. Il sistema di traduzione automatica neurale di Google (GNMT) si basa proprio su questo principio: sfruttare i *Big Data* (si vedano i documenti pubblicati quotidianamente dall'ONU, tradotti in numerose lingue) per confrontare diverse traduzioni in maniera che la rete neurale possa apprendere a tradurre frasi e non parole, quindi a tenere conto del contesto e del significato nel testo. I computer analizzano questi testi cercando schemi statisticamente significativi, cioè schemi tra la traduzione

e il testo originale che è improbabile si verifichino per caso.

Quando il sistema venne lanciato si rivelò sì funzionante, ma ancora colmo di debolezze da raffinare:

- apprendimento della rete lento;
- difficoltà nella traduzione di parole rare;
- esclusione di alcune parole nella traduzione.

Il sistema GNMT

Il sistema si compone, oggi, di due RNNs (*recurrent neural networks*), dotate di *Long Short Term Memories* (LSTM) con 8 strati ciascuna collegati attraverso connessioni residuali. La prima è un codificatore che riceve le frasi da tradurre in input e le codifica in vettori di lunghezza variabile in maniera da farle leggere alla macchina. Alcuni studi [7] [8] (Sutskever et al., 2014; Kalchbrenner and Blunsom, 2013) ritengono che per questo tipo di RNN sia necessario codificare vettori di lunghezza fissa con le parole in input, ma talvolta questo sistema risulta scomodo o addirittura inutile. Il software è una funzione non lineare della forma:

$$\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M = \text{EncoderRNN}(x_1, x_2, \dots, x_M) \quad (1)$$

dove M è il numero di parole di cui è composta la frase in input. Si osservi però che la funzione, prima di produrre $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$, passa da uno stato nascosto che si può indicare con h_t in cui emerge la dipendenza dal tempo t dovuta all'attraversamento della rete neurale *deep* (si approfondirà in seguito la sua struttura). Così la probabilità di ottenere una traduzione \mathbf{Y} data una frase di partenza \mathbf{X} è la probabilità condizionata:

$$P(\mathbf{Y}|\mathbf{X}) = P(\mathbf{Y}|\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M) = \prod_{i=1}^N P(y_i|y_0, y_1, \dots, y_{i-1}; \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M) \quad (2)$$

dove y_0 individua l'inizio della frase (ed è indipendente dalla frase stessa) e \mathbf{Y} è un vettore N -dimensionale con N che è il numero di parole che compongono la frase in output.

Mentre la rete legge i vettori con le parole, si calcola di volta in volta la probabilità per la paro-

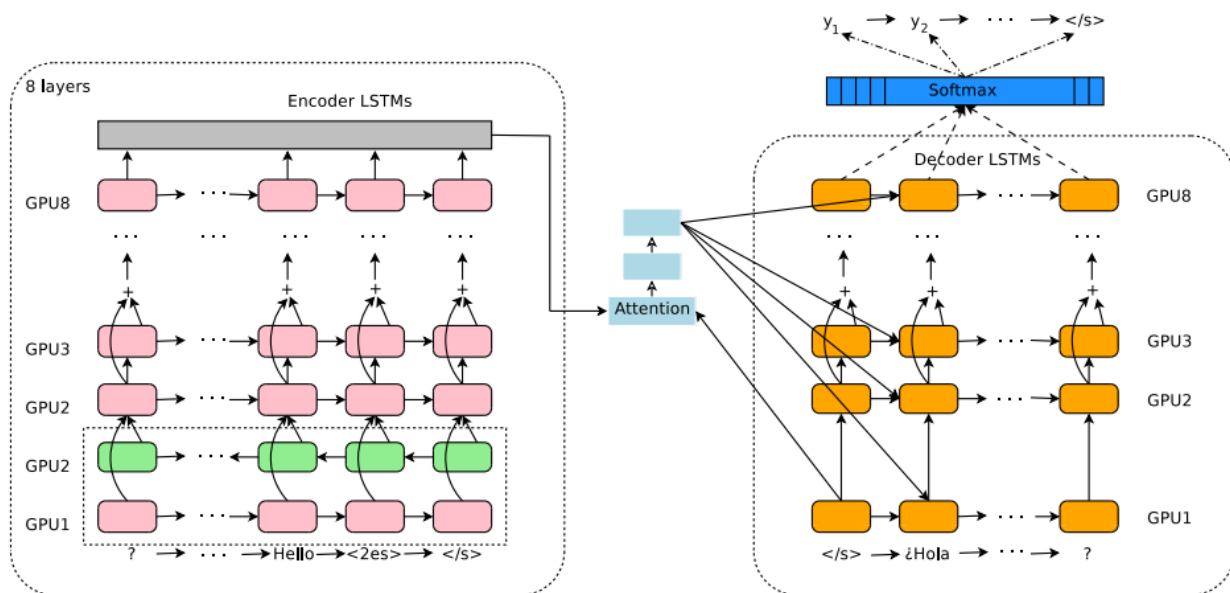


Figura 4: L'architettura del sistema GNMT dotata di un codificatore (a sinistra) con 8 GPU, di cui uno strato bidirezionale, poi un "attention module" e un decodificatore (a destra) con il "softmax layer". Fonte: [6]

la successiva sfruttando la (2). In un primo momento quindi, la rete processa la frase in input attraverso 8 GPU dotate di connessioni residuali tra loro. In particolare uno degli strati è diretto in maniera opposta rispetto agli altri per cogliere meglio il senso della frase che si sta traducendo. La presenza di numerose GPU e di connessioni residuali minimizza la possibilità di dimenticare alcune parole nella traduzione, perché la frase in input viene continuamente richiamata, sia nella fase di apprendimento, sia nella fase di traduzione vera e propria della macchina. Le connessioni inoltre contribuiscono non poco alla velocità di apprendimento della rete, capacità che viene raggiunta anche grazie al parallelismo che caratterizza tutto il modello GNMT.

Una volta che l'informazione viene processata nel codificatore, passa attraverso una struttura di controllo detta "attention module" collegata ad ogni strato del decodificatore (la struttura successiva, anch'essa una RNN). All'interno di quest'ultimo sono presenti sempre 8 GPU che però, questa volta sono dirette tutte nello stesso verso perché qui si opera la traduzione e il contesto è stato già esaminato nella prima parte della rete. Anche qui le GPU sono dotate di connessioni residuali per favorirne la velocità sia in apprendimento sia in azione, e per evitare di tralasciare parole o pezzi di frase utili.

All'interno del decodificatore viene usato un

algoritmo detto di *beam search* per ottenere la frase Y che massimizza la funzione punteggio $s(Y, X)$ in un modello che però sia già allenato a tradurre. Nelle traduzioni si sceglie quindi quella che ottiene il punteggio più alto, infatti la funzione s tiene conto della lunghezza della frase e delle parole, della probabilità calcolata in precedenza e di quanto si stia normalizzando la lunghezza suddetta. Si è osservato che il raddoppiamento del numero di traduzioni analizzate durante il beam search, passando da 4-6 a 8-12, non ha prodotto un grande miglioramento, quindi si valutano a priori dei parametri di controllo che ottimizzano il risultato di $s(Y, X)$. Una volta che la frase viene decodificata negli 8 strati di LSTM, passa nel *softmax layer*, ovvero una funzione esponenziale normalizzata che traduce effettivamente i vettori in parole fino al carattere speciale di fine frase. È affascinante come tutto questo processo, operato da una rete già allenata, si realizzi in un battito di ciglia.

Il wordpiece model

Resta però ancora un problema irrisolto, o apparentemente irrisolto: come fa la rete neurale di Google a tradurre parole rare, o peggio, parole di cui non ha mai visto una traduzione in precedenza? Anche a questo problema, l'algoritmo di Google trova una soluzione molto efficiente:

il cosiddetto *wordpiece model*. Nella codifica delle parole la rete usa suddividere le parole che non conosce in pezzi, per poterli ricercare in altri documenti, testi o parole ed imparare autonomamente il significato di questa parola sconosciuta per poterla tradurre al meglio. Lo stesso principio viene applicato alla traduzione dei nomi propri delle cose, infatti essi non vengono tradotti, perché la rete separa dal resto della parola la prima lettera maiuscola e poi riconosce se sia un nome da tradurre o meno in base al contesto nell'*attention module*. Per far funzionare bene questo tipo di processamento delle parole, Google non solo si serve di vocabolari composti da decine di migliaia di sub-unità di parole, ma in più, assegna ad ogni pezzo di parola nella frase se si tratta della posizione iniziale $\langle b \rangle$, centrale $\langle m \rangle$ o finale $\langle e \rangle$, in maniera da migliorare l'accuratezza nella traduzione.

Limiti e opportunità

Fino a che punto GNMT è in grado di produrre una traduzione *ready-to-use*? Qual è il punto della situazione per la lingua inglese e per le altre lingue? Di seguito si propone a scopo esemplificativo la traduzione in italiano di due frasi dal tedesco (Figura 5) e dall'inglese (Figura 6) svolte da GNMT e da un traduttore umano.

La lingua tedesca possiede una struttura tra le più complesse in riferimento alle lingue germaniche moderne. La frase proposta nell'esempio è stata estratta da un articolo su uno studio psicologico sulla *Schadenfreude*, ovvero la gioia che deriva dalle disgrazie altrui. Osservando le due traduzioni, si può notare come la rete neurale di Google abbia facilmente identificato la struttura della frase, fornendo un output che, dal punto di vista strutturale, potrebbe essere considerato accettabile in base alle convenzioni della lingua di arrivo. Tuttavia, anche un occhio inesperto può rendersi conto di come la frase non abbia senso compiuto, caratteristica fondamentale per una buona comprensione e interpretazione del testo. Nella frase in lingua tedesca sono presenti diversi elementi problematici nel trasferimento da un sistema linguistico all'altro. Primi fra tutti, l'aggettivo *vermeintlich* (presunto), il sostantivo *Überflieger* (persona superiore/brillante) e la locuzione *in den Augen anderer* (agli occhi degli al-

tri) che, trasposti in lingua italiana, creerebbero una ridondanza semantica. Per questo motivo si rende quasi necessaria l'omissione dell'aggettivo e della locuzione, poiché per meglio esprimere il concetto di "persona superiore" occorre esplicitare con mezzi linguistici il fatto che una persona venga considerata superiore da soggetti terzi (sottinteso). La scelta d'uso del plurale nella frase in lingua italiana (nel testo di partenza al singolare) è determinata da convenzioni linguistiche e contestuali per cui, trattandosi di un enunciato generale su un fenomeno sociale, si fa solitamente riferimento ad una categoria di persone e non ad una sola in rappresentanza dell'intero gruppo. Dal punto di vista terminologico, la rete neurale sembra fallire nella disambiguazione dei termini *Überflieger* e *zurückstutzen* che identifica, rispettivamente, come "volantino alto" (traduzione automatica risultante dalla traduzione letterale di *über* e *Flieger*) ed "esecuzione del backup", entrambe fuori dal contesto della frase e, quindi, prive di senso. Un caso particolare è rappresentato dal termine *Schadenfreude*, una delle cosiddette "parole intraducibili" della lingua tedesca. GNMT sembra non dividere il termine in unità linguistiche minime per individuare un equivalente traduttivo e lo restituisce invariato nella traduzione automatica. Si può affermare che i risultati ottenuti da questa simulazione testimoniano un margine di successo ancora molto basso rispetto al livello necessario per una comprensione funzionale del testo e per lo svolgimento di attività come il PEMT.

Se la traduzione dal tedesco all'italiano ha presentato diverse criticità più o meno gravi, diverso è il caso della traduzione dall'inglese all'italiano. La rete neurale di Google fornisce un output di buona qualità, peccando esclusivamente in questioni di carattere terminologico che, ad ogni modo, non costituiscono un ostacolo alla comprensione della frase (vedi *conferenze del partito* vs. *conferenze di partito*). La frase, estratta da un articolo diffuso dallo stesso governo britannico, contiene un termine tecnico in ambito giuridico: *Queen's speech*. Mentre GNMT lo traduce letteralmente, in lingua italiana si rende necessario esplicitare il suo significato per non generare ambiguità e/o fraintendimenti. *Discorso della regina*, traduzione corretta in senso letterale, lascia intendere qualcosa di sostanzialmente diverso

Traduzione DE→IT		
Fonte	„Schadenfreude wirkt dadurch auch sozial regulierend, da sie den vermeintlichen Überflieger in den Augen anderer wieder auf sein menschliches Maß zurückstutzt.“	
GNMT	"Schadenfreude ha quindi un effetto socialmente regolante, in quanto esegue il backup del presunto volantino alto agli occhi degli altri nella sua dimensione umana".	1.0
Traduttore umano	“La gioia maligna agisce, quindi, anche come fattore di riequilibrio sociale, poiché riporta quelli che sono considerati migliori alla loro dimensione più umana.”	4.5

Figura 5: Traduzione della stessa frase DE→IT da "Google Translate" e un traduttore umano con valutazione delle traduzioni su una scala da 0.0 (traduzione errata, inefficace e priva di significato) a 6.0 (traduzione esatta che riporta il significato della frase di partenza)

Traduzione EN→IT		
Fonte	Following the conclusion of the traditional party conference season, the second session of this Parliament will commence with a Queen's Speech on Monday 14 October.	
GNMT	Dopo la conclusione della tradizionale stagione delle conferenze del partito, la seconda sessione di questo Parlamento inizierà con un discorso della regina lunedì 14 ottobre.	5.0
Traduttore umano	Facendo seguito alla chiusura della tradizionale stagione delle conferenze di partito, la seconda sessione di questo Parlamento verrà inaugurata con la lettura del programma del nuovo governo da parte della Regina, il cosiddetto Queen's Speech, lunedì 14 ottobre.	6.0

Figura 6: Traduzione della stessa frase EN→IT da "Google Translate" e un traduttore umano con valutazione delle traduzioni su una scala da 0.0 (traduzione errata, inefficace e priva di significato) a 6.0 (traduzione esatta che riporta il significato della frase di partenza)

rispetto al reale significato, ovvero la *lettura del programma del nuovo governo da parte della Regina*. Tutto sommato, l'output generato dalla rete neurale si dimostra un buon punto di partenza per il passaggio al Post-editing, permettendo una riduzione del tempo tecnico richiesto dall'attività di traduzione.

Conclusioni

I nuovi strumenti a disposizione dei traduttori nell'era dei *Big Data* e della rivoluzione digitale sono risorse da cui non si può più prescindere. La stessa rete neurale di Google si dimostra essere una risorsa affidabile in riferimento alla lingua inglese grazie all'enorme quantità di dati con cui la stessa viene istruita. Allo stesso tempo, questo non può ancora essere garantito per altre lingue, come il tedesco, nel cui caso la traduzione auto-

matica si rivela utile se applicata limitatamente a testi semplificati. Nonostante ciò, i passi da gigante compiuti negli ultimi anni lasciano pensare che si possa presto raggiungere un margine di successo simile a quello per la lingua inglese. La ricerca ha inoltre allargato le sue frontiere spostandosi dalla traduzione scritta alla traduzione istantanea. In seguito al successo delle reti per il riconoscimento vocale, Google punta alla traduzione vocale. Google Translatotron, lanciato a giugno 2019, è un modello sperimentale di traduzione *speech-to-speech* che ha tutto il potenziale di rivoluzionare ancora il mondo digitale e della comunicazione.



- [1] A. DE MAURO, M. GRECO, M. GRIMALDI: "A formal definition of *Big Data* based on its essential features", *Library Review* **65.3** (2016) 122-135.
- [2] R. KITCHIN: "Big Data, new epistemologies and paradigm shifts", *Big Data & society* **1** (2014) .
- [3] K. L. SAKAI: "Language acquisition and brain development", *Science* **310.5749** (2005) 815-819.
- [4] P. K. KUHL: "Brain mechanisms in early language acquisition", *Neuron* **67.5** (2010) 715-716.
- [5] P. CHERUBINI, V. ZURLONI: *Psicologia cognitiva. Mente e cervello. Con e-text. Con espansione online.* , (2014).447-489
- [6] W. YONGHUI, M. SCHUSTER, C. ZHIFENG, Q. V. LE, M. NOROUZI: "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation", (2016) .
- [7] K. CHO, B. VAN MERRIËNBOER, D. BAHDANAU, Y. BENGIO: "On the properties of neural machine translation: Encoder-Decoder approaches.", *Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*. (2014) .
- [8] N. KALCHBRENNER, P. BLUNSOM: "Recurrent continuous translation models", *Proceedings of the ACL Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics. (2013) .



Caterina Milone: Caterina Milone è una studentessa ISUFI del IV anno dell'area di Scienze Sociali e studia Metodologia dell'intervento psicologico presso l'Università del Salento dopo aver conseguito la laurea triennale in Scienza e tecniche psicologiche con Lode nello stesso ateneo.

Aurora Paladini: Aurora Paladini è una studentessa ISUFI del IV anno dell'area di Scienze Umane e ha conseguito la laurea triennale in Scienza e tecniche della Mediazione Linguistica con Lode presso l'Università del Salento.

Lorenzo Pellegrino: Lorenzo Pellegrino è uno studente ISUFI al III anno dell'area di Scienze Naturali e frequenta il III anno del Corso di Laurea in Matematica presso il Dipartimento di Matematica e Fisica "Ennio De Giorgi" dell'Università del Salento.