



Lost in Translation

Giuseppe Attardi
Università di Pisa

Machine Translation, circa 2001

Lo spirito è forte ma la carne è debole

tradotto in russo

La vodka è forte ma la bistecca è tenera

apocrifo

I sinonimi di Altavista, Umberto Eco. 2007

Babelfish 2004

Originale inglese	Traduzione italiana
The Works of Shakespeare	Gli impianti di Shakespeare
Harcourt Brace	sostegno di Harcourt
Speaker of the chamber of deputies	Altoparlante dell'alloggiamento dei delegati
Studies in the logic of Charles Sanders Peirce	Studi nella logica delle sabbiatrici Peirce del Charles

Google 2007

Originale inglese	Traduzione italiana
The Works of Shakespeare	Le opere di Shakespeare
Harcourt Brace	Harcourt Brace
Speaker of the chamber of deputies	Presidente della Camera dei deputati
Studies in the logic of Charles Sanders Peirce	Studi nella logica di Charles Sanders Peirce

Progressi nella Machine Translation

- Gli chiese di riorganizzare Forza Italia
The churches to reorganize Italy Force (Altavista)
She asked him to reorganize Forza Italia (Google)
- Il ministro Stanca si è laureato alla Bocconi
The Minister Stanca graduated at Mouthfuls (Altavista)
The Minister Stanca is a graduate of Bocconi (Google)

How to learn **natural** language

- Children learn to speak **naturally**, by interacting with adults
- Nobody teaches them grammar
- Is it possible to let computer learn language in a similarly **natural** way?

Statistical Machine Learning

- Supervised Training
- Annotated document collections
- Ability to process Big Data
 - If we used same algorithms 10 years ago they would still be running
- Similar techniques for speech and text

Machine Translation

黛玉自在枕上感念宝钗。。。又听见窗外竹梢蕉叶之上，雨声淅沥，清寒透幕，不觉又滴下泪来。

The Story of the Stone (“The Dream of the Red Chamber”)

- Cao Xueqin 1792

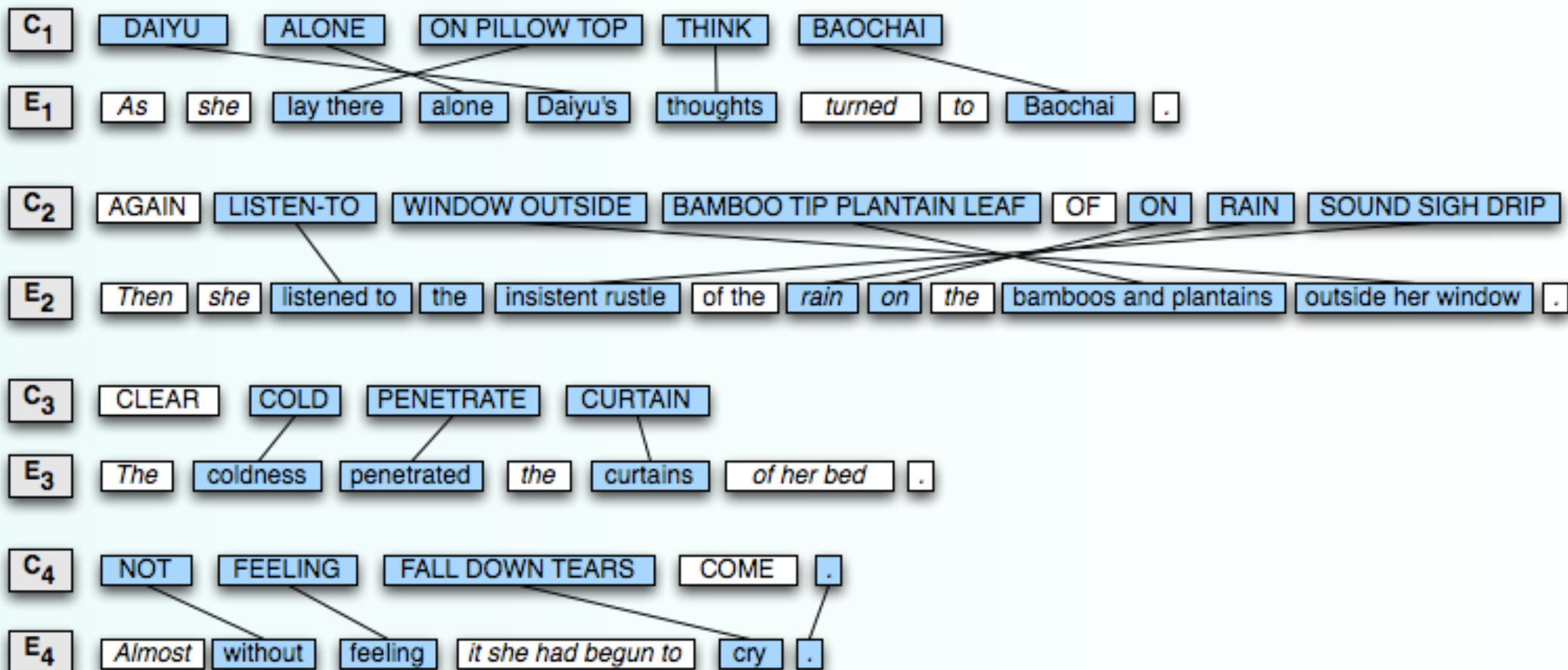
Chinese gloss: Dai-yu alone on bed top think-of-with-gratitude Bao-chai again listen to window outside bamboo tip plantain leaf of on-top rain sound sigh drop clear cold penetrate curtain not feeling again fall down tears come

Hawkes translation: As she lay there alone, Dai-yu’s thoughts turned to Bao-chai. Then she listened to the insistent rustle of the rain on the bamboos and plantains outside her window. The coldness penetrated the curtains of her bed. Almost without noticing it she had begun to cry.

Machine Translation

- Difficoltà:
 - Suddivisione in frasi: 4 in inglese, 1 in cinese
 - Differenze grammaticali
 - Il cinese non riporta il tempo:
 - As, turned to, had begun
 - *tou* ⇒ penetrated
 - Né pronomi né articoli in cinese
 - Differenze stilistiche e culturali differences
 - Bamboo tip plantain leaf ⇒ bamboos and plantains
 - Ma 'curtain' ⇒ curtains of her bed
 - Rain sound sigh drop ⇒ insistent rustle of the rain

Alignment in Machine Translation



Warren Weaver (1947)

When I look at an article in Russian, I say to myself: This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode.



Warren Weaver memo

- <http://www.stanford.edu/class/linguist289/weaver001.pdf>
- “There are certain invariant properties which are... to some statistically useful degree, common to all languages.”
- On March 4, 1947, “having considerable exposure to computer design problems during the war, and being aware of the speed, capacity, and logical flexibility possible in modern electronic computers”, Weaver suggested that computers to be used for translation

Rosetta Stone

- Carved in 196 BC
- Found in 1799
- Decoded in 1822



Egyptian hieroglyphs

Egyptian Demotic

Greek

Centauri/Arcturan

Come tradurre questo in Arcturan:

farok crrrok hihok yorok klok kantok ok-yurp

Centauri/Arcturan Parallel Corpus

1a. ok-voon ororok sprok . 1b. at-voon bichat dat .	7a. lalok farok ororok lalok sprok izok enemok . 7b. wat jjat bichat wat dat vat eneat .
2a. ok-drubel ok-voon anak plok sprok . 2b. at-drubel at-voon pippat rrat dat .	8a. lalok brok anak plok nok . 8b. iat lat pippat rrat nnat .
3a. erok sprok izok hihok ghrok . 3b. totat dat arrat vat hilat .	9a. wiwok nok izok kantok ok-yurp . 9b. totat nnat quat oloat at-yurp .
4a. ok-voon anak drok brok jok . 4b. at-voon krat pippat sat lat .	10a. lalok mok nok yorok ghrok klok . 10b. wat nnat gat mat bat hilat .
5a. wiwok farok izok stok . 5b. totat jjat quat cat .	11a. lalok nok crrrok hihok yorok zanzanok . 11b. wat nnat arrat mat zanzanat .
6a. lalok sprok izok jok stok . 6b. wat dat krat quat cat .	12a. lalok rarok nok izok hihok mok . 12b. wat nnat forat arrat vat gat .

Traduzione da Centauri ad Arcturan

Tradurre in Arcturan: **farok** crrrok hihok yorok klok kantok ok-yurp

ok-voon ororok sprok . at-voon bichat dat .	lalok farok ororok lalok sprok izok enemok . wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anak plok sprok . at-drubel at-voon pippat rrat dat .	lalok brok anak plok nok . iat lat pippat rrat nnat .
erok sprok izok hihok ghrok . totat dat arrat vat hilat .	wiwok nok izok kantok ok-yurp . totat nnat quat oloat at-yurp .
ok-voon anak drok brok jok . at-voon krat pippat sat lat .	lalok mok nok yorok ghrok klok . wat nnat gat mat bat hilat .
wiwok farok izok stok . totat jjat quat cat .	lalok nok crrrok hihok yorok zanzanok . wat nnat arrat mat zanzanat .
lalok sprok izok jok stok . wat dat krat quat cat .	lalok rarok nok izok hihok mok . wat nnat forat arrat vat gat .

Traduzione da Centauri ad Arcturan

ok-voon ororok sprok . at-voon bichat dat .	lalok farok ororok lalok sprok izok enemok . wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anak plok sprok . at-drubel at-voon pippat rrat dat .	lalok brok anak plok nok . iat lat pippat rrat nnat .
erok sprok izok hihok ghirok . totat dat arrat vat hilat .	wiwok nok izok kantok ok-yurp . totat nnat quat oloat at-yurp .
ok-voon anak drok brok jok . at-voon krat pippat sat lat .	lalok mok nok yorok ghirok klok . wat nnat gat mat bat hilat .
wiwok farok izok stok . totat jjat quat cat .	lalok nok crrok hihok yorok zanzanok . wat nnat arrat mat zanzanat .
lalok sprok izok jok stok . wat dat krat quat cat .	lalok rarok nok izok hihok mok . wat nnat forat arrat vat gat .

Allineamento tra parole Centauri/Arcturan

Tradurre in Arcturan: **farok** **crrok** hihok yorok klok kantok ok-yurp

1a. ok-voon ororok sprok .

(

1b. at-voon bichat dat .

2a. ok-drubel ok-voon anak plok sprok .

2b. at-drubel at-voon pippat rrat dat .

3a. erok sprok izok hihok ghirok .

3b. totat dat arrat vat hilat .

4a. ok-voon anak drok brok jok .

4b. at-voon krat pippat sat lat .

5a. wiwok **farok** izok stok .

/

5b. totat jjat quat cat .

6a. lalok sprok izok jok stok .

6b. wat dat krat quat cat .

7a. lalok **farok** ororok lalok sprok izok enemok .

/

7b. wat jjat bichat wat dat vat eneat .

8a. lalok brok anak plok nok .

8b. iat lat pippat rrat nnat .

9a. wiwok nok izok kantok ok-yurp .

9b. totat nnat quat oloat at-yurp .

10a. lalok mok nok yorok ghirok klok .

10b. wat nnat gat mat bat hilat .

11a. lalok nok **crrok** hihok yorok zanzanok .

???

11b. wat nnat arrat mat zanzanat .

12a. lalok rarok nok izok hihok mok .

12b. wat nnat forat arrat vat gat .

Centauri/Arcturan Alignment

Your assignment, translate this to Arcturan: **farok** crrrok **hihok yorok** **clock** kantok ok-yurp

1a. ok-voon ororok sprok .

1b. at-voon bichat dat .

2a. ok-drubel ok-voon anak plok sprok .

2b. at-drubel at-voon pippat rrat dat .

3a. erok sprok izok hihok ghirok .

3b. totat dat arrat vat hilat .

4a. ok-voon anak drok brok jok .

4b. at-voon krat pippat sat lat .

5a. wiwok farok izok stok .

5b. totat jjat quat cat .

6a. lalok sprok izok jok stok .

6b. wat dat krat quat cat .

7a. lalok farok ororok lalok sprok izok enemok .

7b. wat jjat bichat wat dat vat eneat .

8a. lalok brok anak plok nok .

8b. iat lat pippat rrat nnat .

9a. wiwok nok izok kantok ok-yurp .

9b. totat nnat quat oloat at-yurp .

10a. lalok mok nok yorok ghirok **clock** .

10b. wat nnat gat mat bat hilat .

11a. lalok nok crrrok hihok yorok zanzanok .

11b. wat nnat arrat mat zanzanat .

12a. lalok rarok nok izok hihok mok .

12b. wat nnat forat arrat vat gat .

process of
elimination

Centauri/Arcturan Alignment

Your assignment, translate this to Arcturan: **farok** **errrok** **hihok** **yorok** **clok** kantok ok-yurp

1a. ok-voon ororok sprok .

1b. at-voon bichat dat .

2a. ok-drubel ok-voon anak plok sprok .

2b. at-drubel at-voon pippat rrat dat .

3a. erok sprok izok hihok ghirok .

3b. totat dat arrat vat hilat .

4a. ok-voon anak drok brok jok .

4b. at-voon krat pippat sat lat .

5a. wiwok farok izok stok .

5b. totat jjat quat cat .

6a. lalok sprok izok jok stok .

6b. wat dat krat quat cat .

7a. lalok farok ororok lalok sprok izok enemok .

7b. wat jjat bichat wat dat vat eneat .

8a. lalok brok anak plok nok .

8b. iat lat pippat rrat nnat .

9a. wiwok nok izok kantok ok-yurp .

9b. totat nnat quat oloat at-yurp .

10a. lalok mok nok yorok ghirok klok .

10b. wat nnat gat mat bat hilat .

11a. lalok nok **errrok** hihok yorok **zanzanok** .

11b. wat nnat arrat mat zanzanat .

12a. lalok rarok nok izok hihok mok .

12b. wat nnat forat arrat vat gat .

cognate?

Traduzione da Centauri ad Arcturan

ok-voon ororok sprok . at-voon bichat dat .	lalok farok ororok lalok sprok izok enemok . / wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anak plok sprok . at-drubel at-voon pippat rrat dat .	lalok brok anak plok nok . / iat lat pippat rrat nnat .
erok sprok izok hihok ghrok . / / totat dat arrat vat hilat .	wiwok nok izok kantok ok-yurp . totat nnat quat oloat at-yurp .
ok-voon anak drok brok jok . at-voon krat pippat sat lat .	lalok mok nok yorok ghrok klok . X / / wat nnat gat mat bat hilat .
wiwok farok izok stok . / totat jjat quat cat .	lalok nok crrok hihok yorok zanzanok . / ← / / wat nnat arrat mat zanzanat .
lalok sprok izok jok stok . wat dat krat quat cat .	lalok rarok nok izok hihok mok . / / / wat nnat forat arrat vat gat .

zero
fertility

Centauri/Arcturan Alignment

Your assignment, put these words in order:

{ jjat, arrat, mat, bat, oloat, at-yurp }

1a. ok-voon ororok sprok .

1b. at-voon bichat dat .

2a. ok-drubel ok-voon anak plok sprok .

2b. at-drubel at-voon pippat rrat dat .

3a. erok sprok izok hihok ghirok .

3b. totat dat arrat vat hilat .

4a. ok-voon anak drok brok jok .

4b. at-voon krat pippat sat lat .

5a. wiwok farok izok stok .

5b. totat jjat quat cat .

6a. lalok sprok izok jok stok .

6b. wat dat krat quat cat .

7a. lalok farok ororok lalok sprok izok enemok .

7b. wat jjat bichat wat dat vat eneat .

8a. lalok brok anak plok nok .

8b. iat lat pippat rrat nnat .

9a. wiwok nok izok kantok ok-yurp .

9b. totat nnat quat oloat at-yurp .

10a. lalok mok nok yorok ghirok klok .

10b. wat nnat gat mat bat hilat .

11a. lalok nok **crrok** hihok yorok zanzanok .

11b. wat nnat arrat mat zanzanat .

12a. lalok rarok nok izok hihok mok .

12b. wat nnat forat arrat vat gat .

zero
fertility

In Realtà sono Spagnolo/Inglese

Clients do not sell pharmaceuticals in Europe => Clientes no venden medicinas en Europa

Garcia and associates .
Garcia y asociados .

the clients and the associates are enemies .
los clientes y los asociados son enemigos .

Carlos Garcia has three associates .
Carlos Garcia tiene tres asociados .

the company has three groups .
la empresa tiene tres grupos .

his associates are not strong .
sus asociados no son fuertes .

its groups are in Europe .
sus grupos estan en Europa .

Garcia has a company also .
Garcia tambien tiene una empresa .

the modern groups sell strong pharmaceuticals .
los grupos modernos venden medicinas fuertes .

its clients are angry .
sus clientes estan enfadados .

the groups do not sell zenzanine .
los grupos no venden zanzanina .

the associates are also angry .
los asociados tambien estan enfadados

the small groups are not modern .
los grupos pequenos no son modernos .

.

History of MT: Pessimism

- 1959/1960: Bar-Hillel “Report on the state of MT in US and GB”
 - Argued FAHQT too hard (semantic ambiguity, etc.)
 - Should work on semi-automatic instead of automatic
 - His argument:
Little John was looking for his toy box. Finally, he found it. The box was in the pen. John was very happy.
 - Only human knowledge lets us know that ‘playpens’ are bigger than boxes, but ‘writing pens’ are smaller
 - His claim: we would have to encode all of human knowledge

Breve Storia della MT

1946 Booth e Weaver ne discutono alla Rockefeller Foundation

1947-48 idea di fare traduzione diretta basata su dictionari

1949 Weaver memorandum diffonde l'idea

1952 tutti i 18 ricercatori al mondo in MT si incontrano al MIT

1954 IBM/Georgetown Demo Russian-English MT

1955-65 molti laboratori iniziano a lavorare su MT

Inizi della MT: 1950s

- Early NLP (Machine Translation) on machines less powerful than pocket calculators
- Foundational work on automata, formal languages, probabilities, and information theory
- First speech systems (Davis et al., Bell Labs)
- MT heavily funded by military – a lot of it was just word substitution programs but there were a few seeds of later successes, e.g., trigrams
- Little understanding of natural language syntax, semantics, pragmatics
- Problem soon appeared intractable

Pessimismo

- The ALPAC report
 - Coordinato da John R. Pierce of Bell Labs
 - Conclusioni:
 - Supply of human translators exceeds demand
 - All the Soviet literature is already being translated
 - MT has been a failure: all current MT work had to be post-edited
 - Sponsored evaluations which showed that intelligibility and informativeness was worse than human translations
 - Effetti:
 - MT research suffered
 - Funding loss
 - Number of research labs declined
 - Association for Machine Translation and Computational Linguistics dropped MT from its name

Risorgimento

1976 Meteo, weather forecasts from English to French

1970's

European focus in MT; mainly ignored in US

1980's

ideas of using early AI techniques in MT (KBMT, CMU)

Focus on "interlingua" systems, especially in Japan

1990's

Commercial MT systems

Statistical MT

Speech-to-speech translation

2000's

Statistical MT takes off

Google Translate

1990s-2010s: Statistical Machine Translation

- SMT is a huge research field
 - The best systems are extremely complex
 - Hundreds of important details, too many to mention
 - Systems have many separately-designed subcomponents
 - Lots of feature engineering
 - Need to design features to capture particular language phenomena
- Require compiling and maintaining extra resources
 - Like tables of equivalent phrases
- Lots of human effort to maintain
 - Repeated effort for each language pair!

2014: Neural Machine Translation

Impatto enorme sulla ricerca in
Machine Translation

NMT: la storia di maggior successo del Deep Learning

Neural Machine Translation è passata da attività di ricerca marginale nel **2014** a metodo principale nel **2016**

- **2014:** Primo articolo pubblicato sul seq2seq (trasformazione di sequenze in sequenze)
- **2016:** Google Translate passa da SMT a NMT
- Stupefacente!
 - Sistemi di **SMT**, costruiti da centinaia di informatici lungo periodi di anni, surclassati da sistemi di NMT allenati da una manciata di sviluppatori in pochi mesi

Tsunami Tecnologico

- Mai tecnologia linguistica è passata così rapidamente da ricerca a utilizzo commerciale

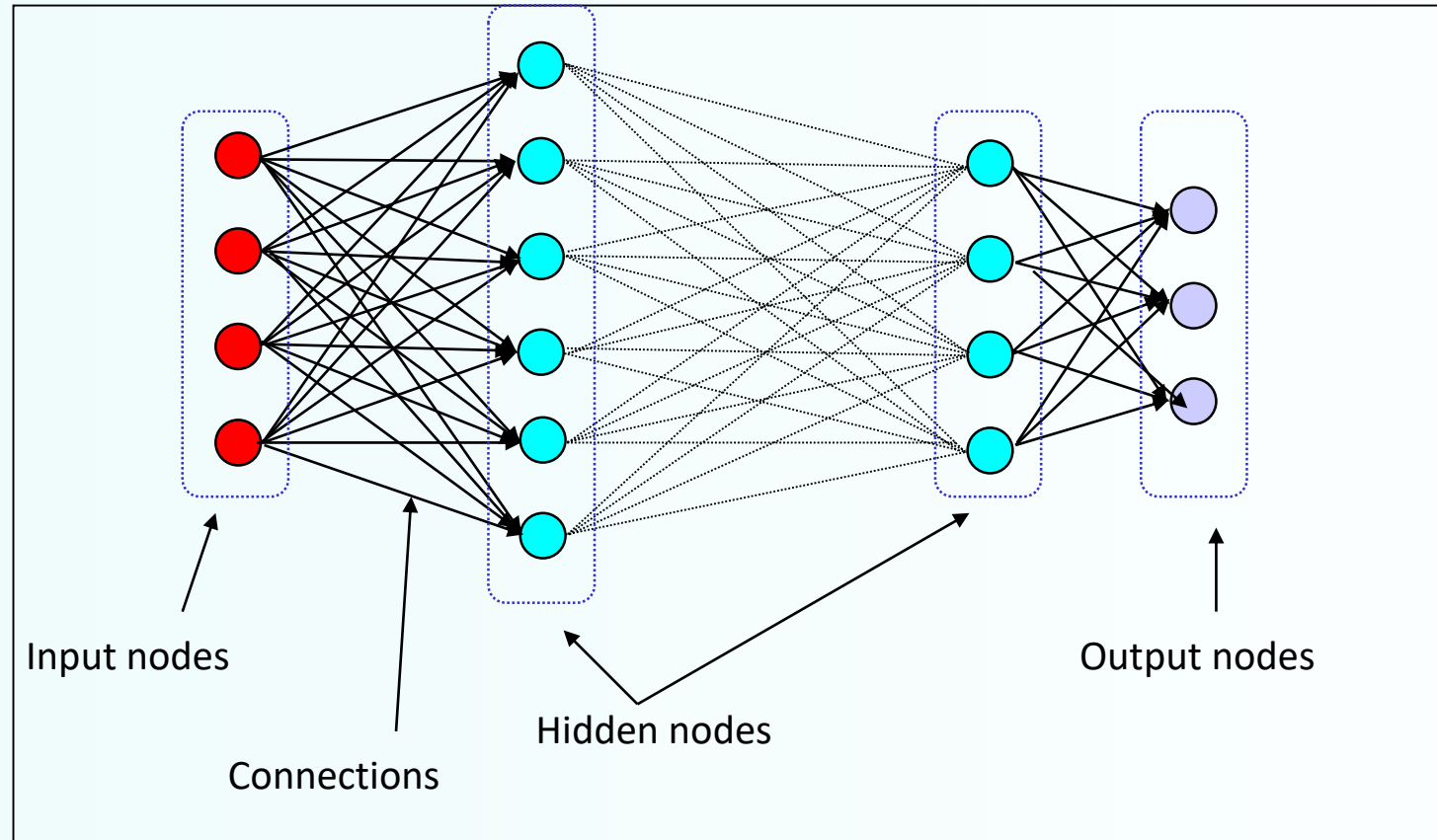
9/2016 Google Brain annuncia traduttore NMT

10/2016 Systran segue

11/2016 Microsoft fa altrettanto

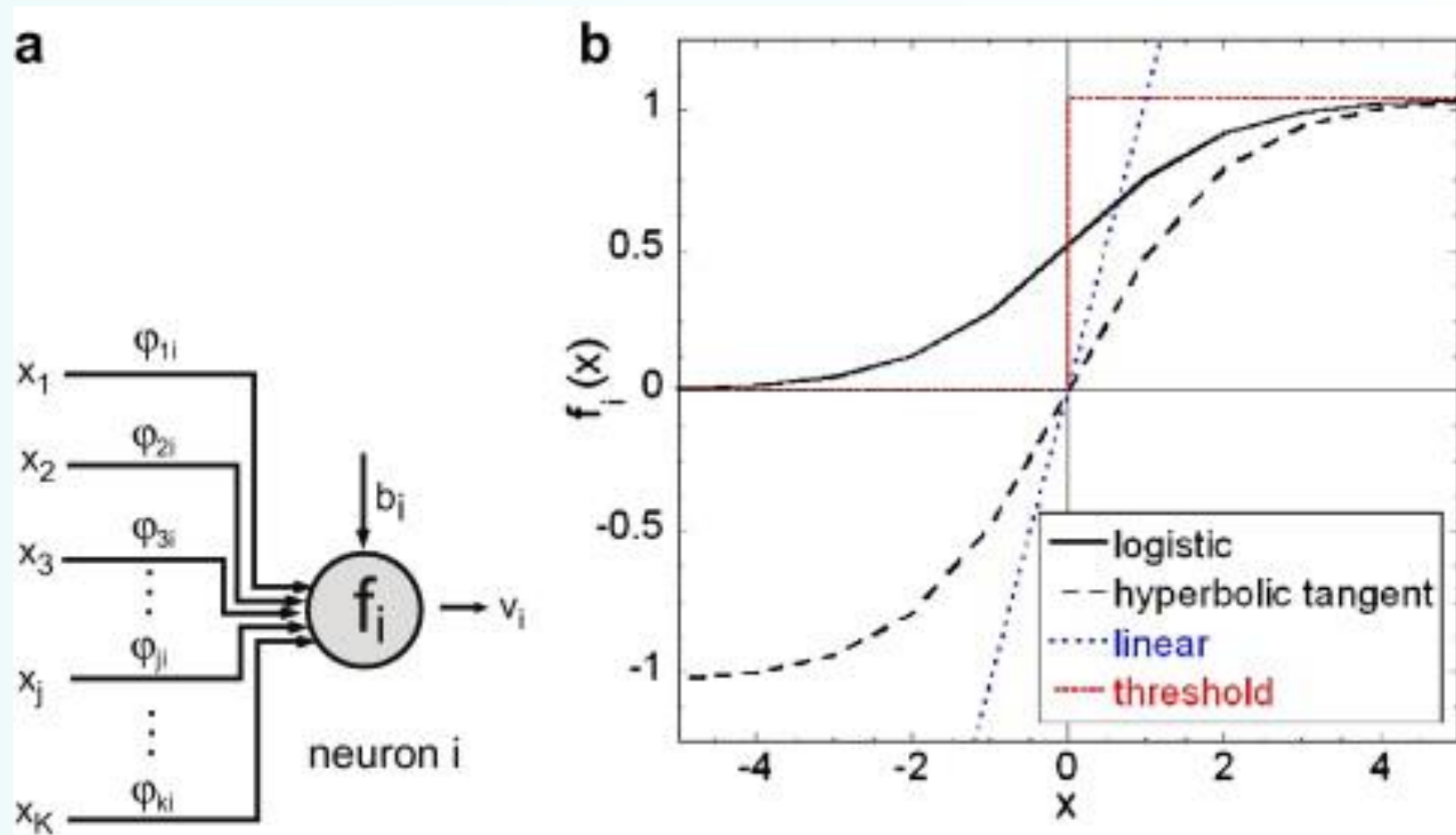
6/2018 Studenti del mio corso di NLP in grado di sviluppare traduttore inglese/italiano in 3 settimane

Deep Neural Network



$$\begin{aligned} \text{Output : } y_i &= f(w_i^1 x_1 + w_i^2 x_2 + w_i^3 x_3 + \Lambda + w_i^m x_m) \\ &= f\left(\sum_j w_i^j x_j\right) \end{aligned}$$

Activation Functions



Apprendimento: Backpropagation Gradient Descent

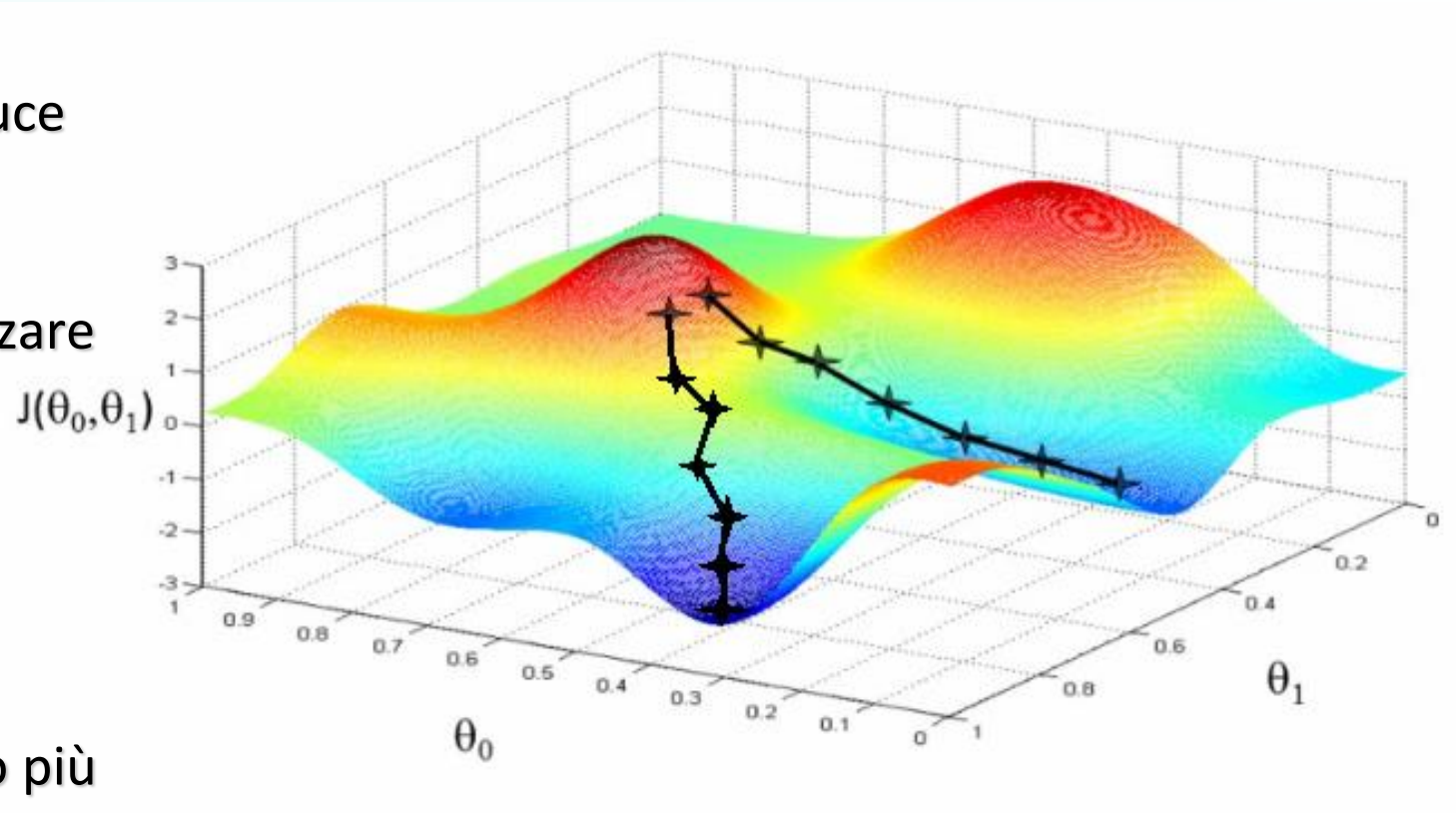
Dati esempi di apprendimento, ossia coppie <input, output>

La rete, applicata all'input, produce una previsione di output

Per apprendere bisogna minimizzare l'errore tra previsione e output

Si parte dal punto indicato come errore sulla superficie

Si cerca di scendere verso il punto più basso



Idea antica

- Introdotta negli anni '60
- Tentata per decenni da anni '80, Rumelhart, Hinton and Williams
- Abbandonata perché non funzionava abbastanza bene

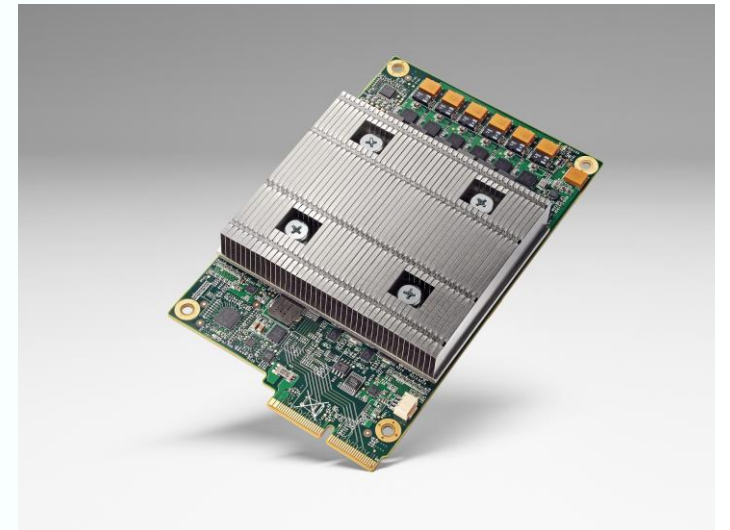
Deep Learning Breakthrough: 2006

- Hinton, Osindero & Teh. 2016.
- Bengio, Lamblin, Popovici, Larochelle. 2006.
- Ranzato, Poultney, Chopra, LeCun. 2006.
- LeCun racconta che il suo articolo sulle Convolutional Neural Network venne rifiutato nel 2011 a NIPS.
L'anno successivo non c'era articolo al NIPS che non usasse le CNN



Breakthrough Tecnologici

- Migliori tecniche di Machine learning
- Grandi quantità di dati
- Notevoli capacità di calcolo
 - Acelaratori grafici (GPU)
 - O dedicati: Tensor Processig Unit



AlphaGo

- TPU can fit into a hard drive slot within the data center rack and has already been powering RankBrain and Street View



Unreasonable Effectiveness of Data

- Halevy, Norvig, and Pereira argue that we should stop acting as if our goal is to author extremely elegant theories, and instead embrace complexity and make use of the best ally we have: the unreasonable effectiveness of data.
- A simpler technique on more data beat a more sophisticated technique on less data.
- Language in the wild, just like human behavior in general, is messy.

Scientific Dispute: is it science?



Prof. Noam Chomsky, Linguist, MIT



Peter Norvig, Director of research, Google

Neural Machine Translation

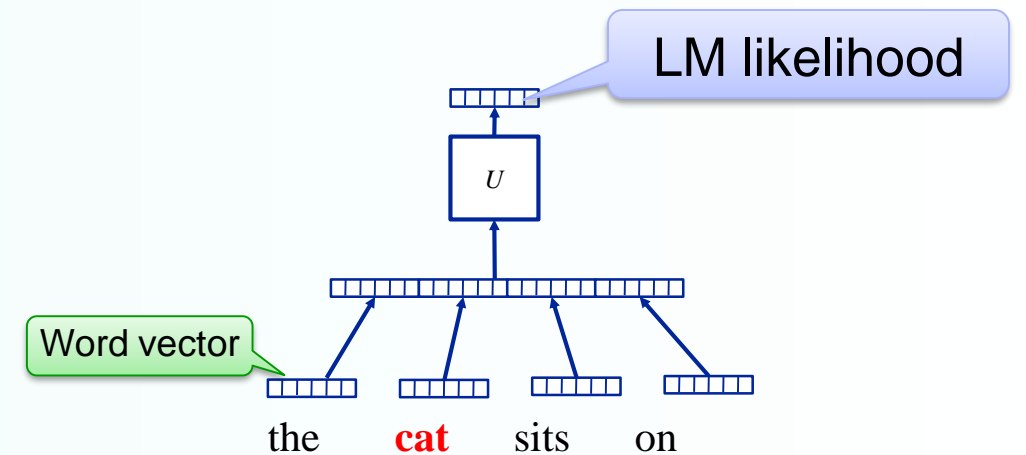
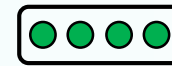
- **Neural Machine Translation** (NMT) utilizza una *singola Rete Neurale* per tradurre una frase in un'altra lingua
- L'apprendimento avviene a partire da un *largo corpus parallelo*

E NIENTE ALTRO!!!

- Niente dizionari, niente grammatiche, niente regole,
- Il processo si chiama **sequence-to-sequence** (aka **seq2seq**) e coinvolge *due RNN* (Recurrent Neural Network)

E il senso delle Parole?

- Word Embeddings
- A ogni parola viene associato un vettore di pesi che rappresenta centinaia di sottili aspetti che rappresentano aspetti del significato della parola
- Come si calcolano?
- Automaticamente con una Rete Neurale che impara a conoscere il senso della parole da una enorme raccolta di testi



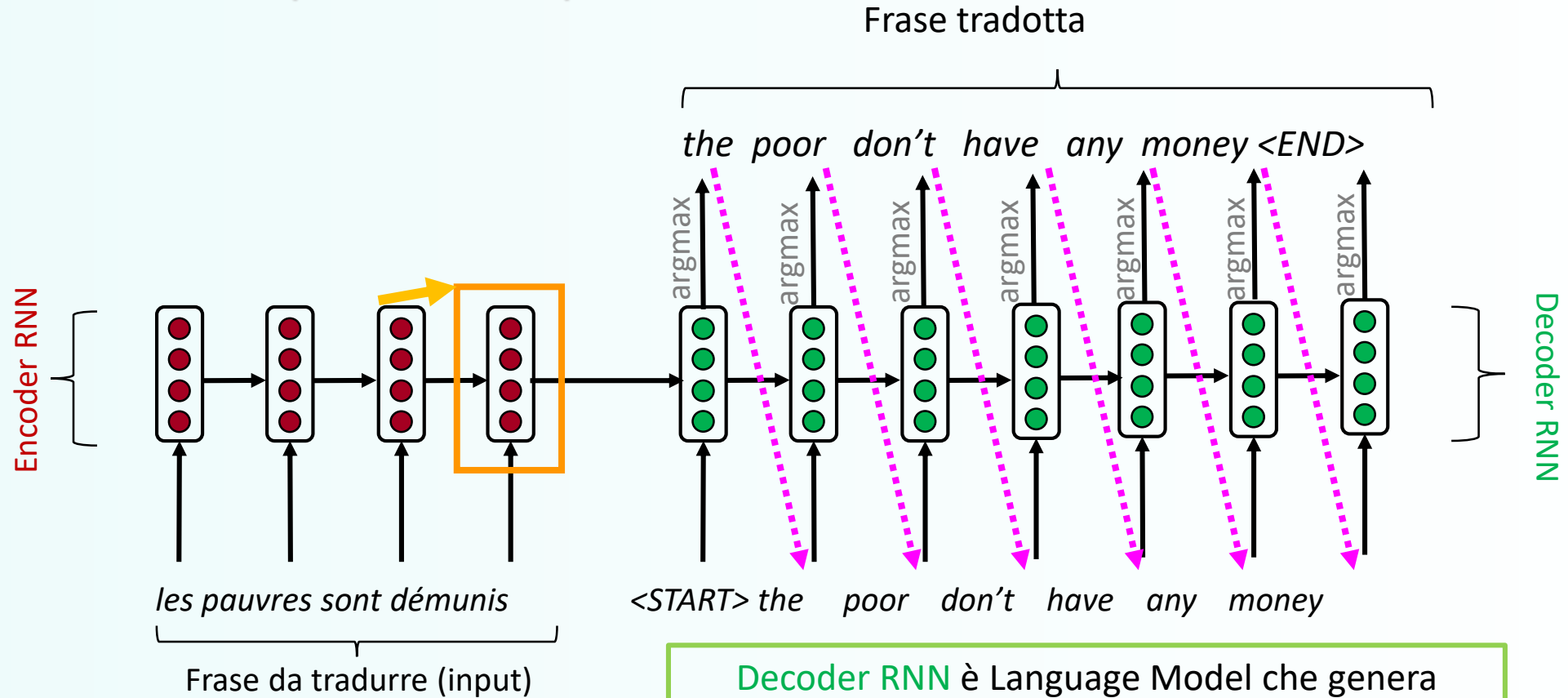
Word Embeddings

FRANCE	JESUS	XBOX	REDDISH	SCRATCHED	MEGABITS
454	1973	6909	11724	29869	87025
AUSTRIA	GOD	AMIGA	GREENISH	NAILED	OCTETS
BELGIUM	SATI	PLAYSTATION	BLUISH	SMASHED	MB/S
GERMANY	CHRIST	MSX	PINKISH	PUNCHED	BIT/S
ITALY	SATAN	IPOD	PURPLISH	POPPED	BAUD
GREECE	KALI	SEGA	BROWNISH	CRIMPED	CARATS
SWEDEN	INDRA	PSNUMBER	GREYISH	SCRAPED	KBIT/S
NORWAY	VISHNU	HD	GRAYISH	SCREWED	MEGAHERTZ
EUROPE	ANANDA	DREAMCAST	WHITISH	SECTIONED	MEGAPIXELS
HUNGARY	PARVATI	GEFORCE	SILVERY	SLASHED	GBIT/S
SWITZERLAND	GRACE	CAPCOM	YELLOWISH	RIPPED	AMPERES

neighboring words **are** semantically related

Neural Machine Translation (NMT)

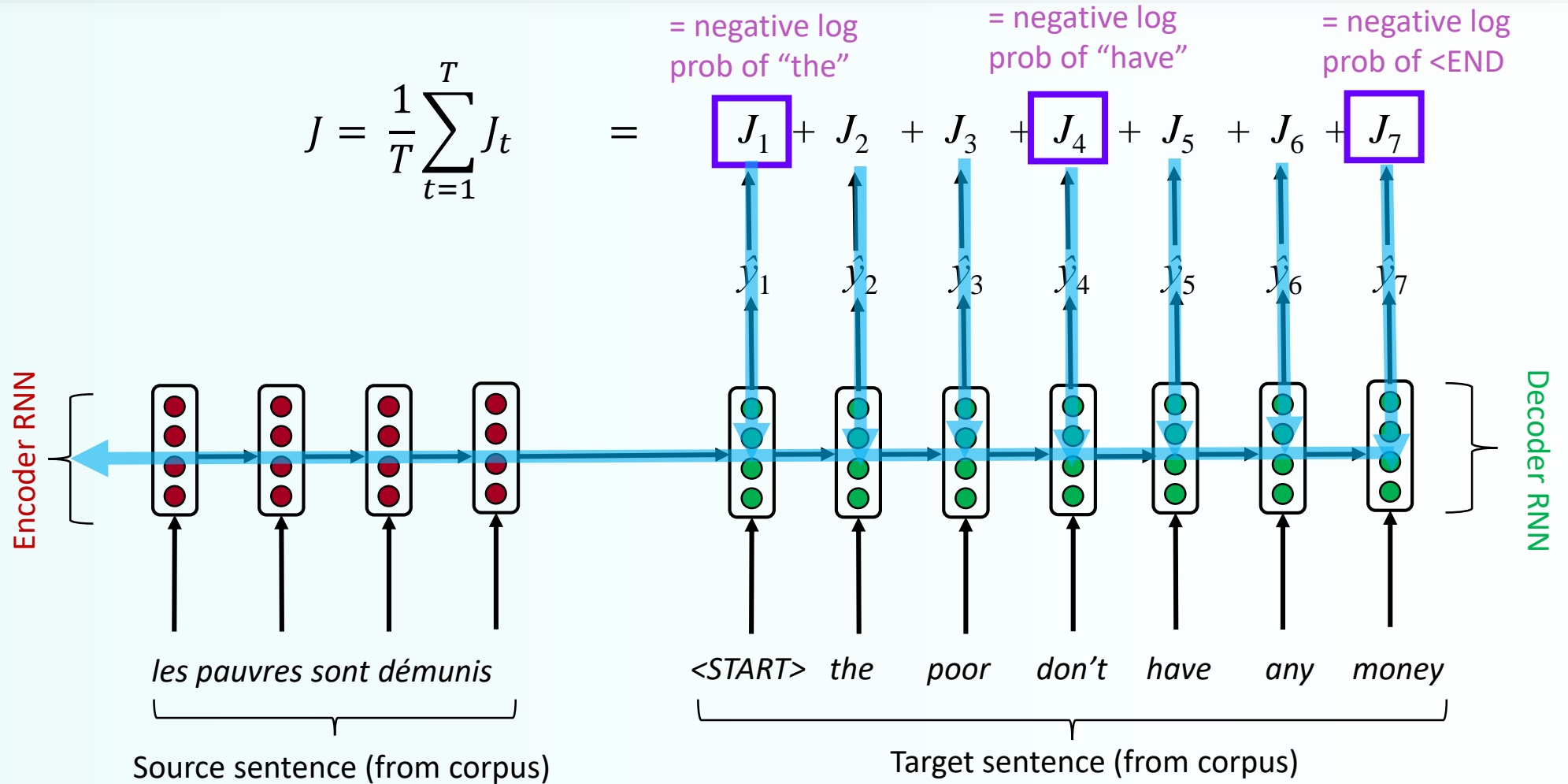
Il modello sequence to sequence



Encoder RNN produce una **codifica** della frase da tradurre

Decoder RNN è Language Model che genera la frase tradotta sulla base dell'**encoding**.

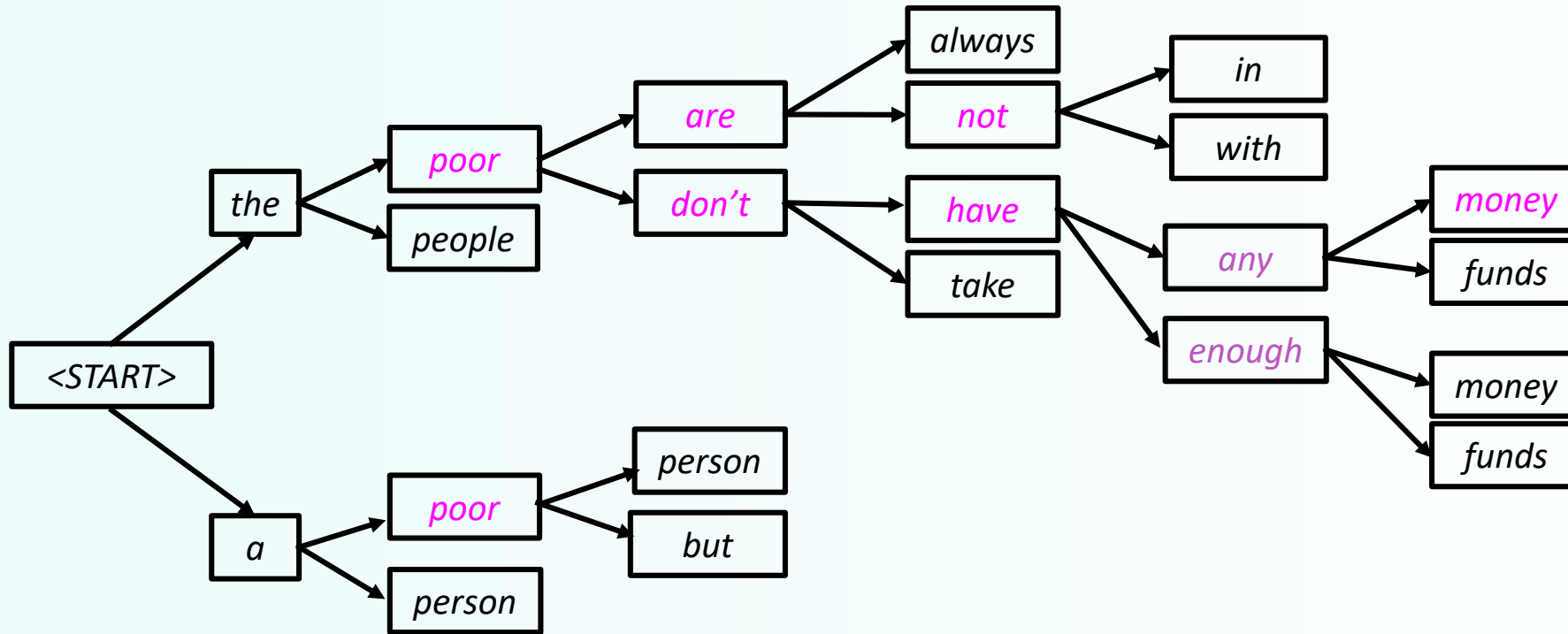
Training a Neural Machine Translation system



Seq2seq is optimized as a **single system**.
Backpropagation operates "end to end"

Non sempre la prima scelta è la migliore

Beam size = 2



Pro/Cons di NMT rispetto a SMT

Benefici

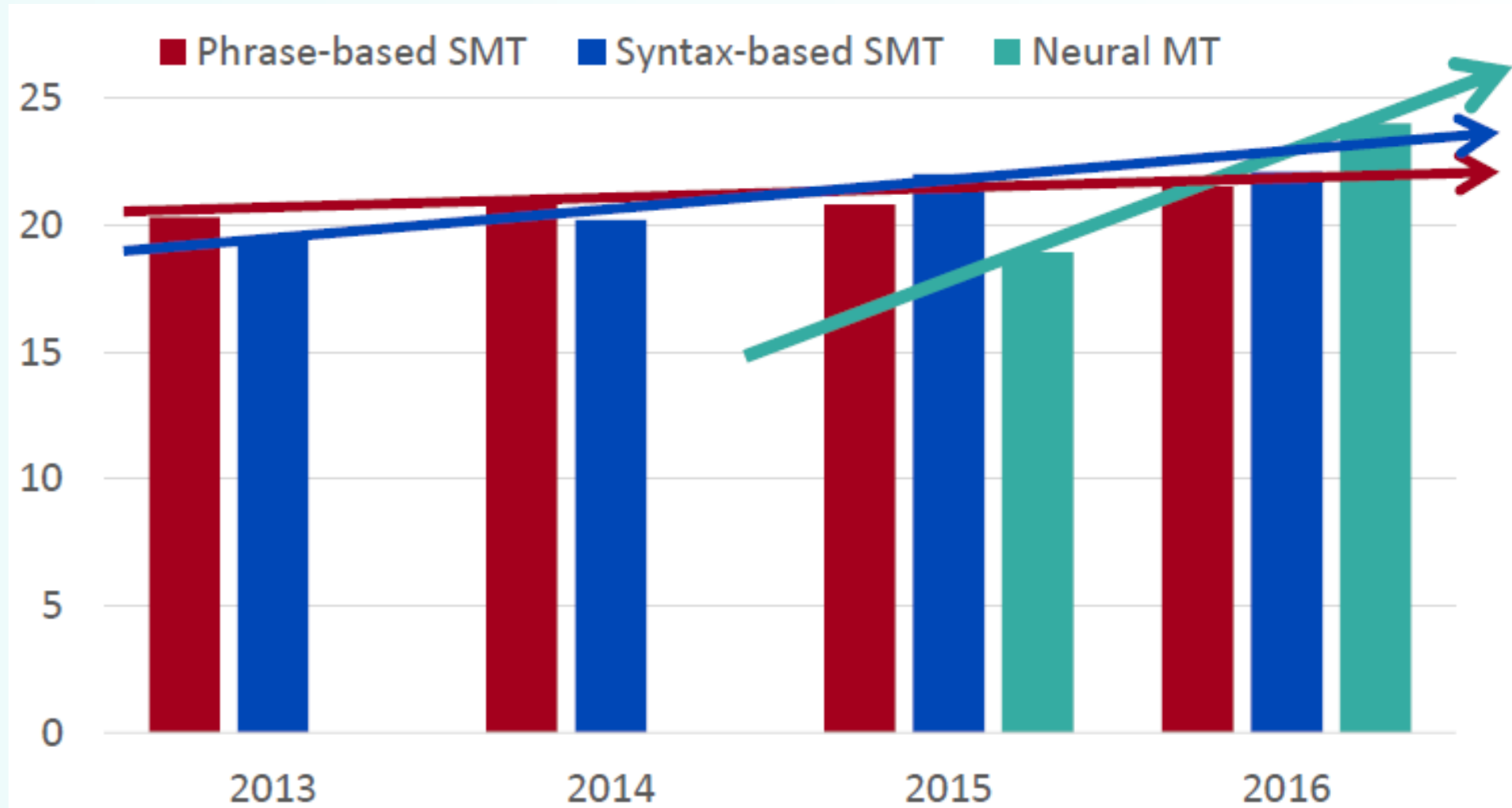
- Migliori prestazioni
 - più **fluente**
 - Miglior uso del **contesto**
 - Miglior uso delle **somiglianze tra frasi**
- Una **sola rete neural** che fa tutto il lavoro end-to-end
- Richiede molto **meno sforzo umano di ingegnerizzazione**
 - Elimina faticosa messa a punto
 - Funziona per qualunque coppia di linguaggi

Svantaggi

- NMT è **meno interpretabile**
 - Difficile da debuggare
- NMT è **difficile da controllare**
 - Per esempio, non è agevole specificare regole o guide per la traduzione
 - Preoccupazioni etiche!

Progresso nel tempo della MT

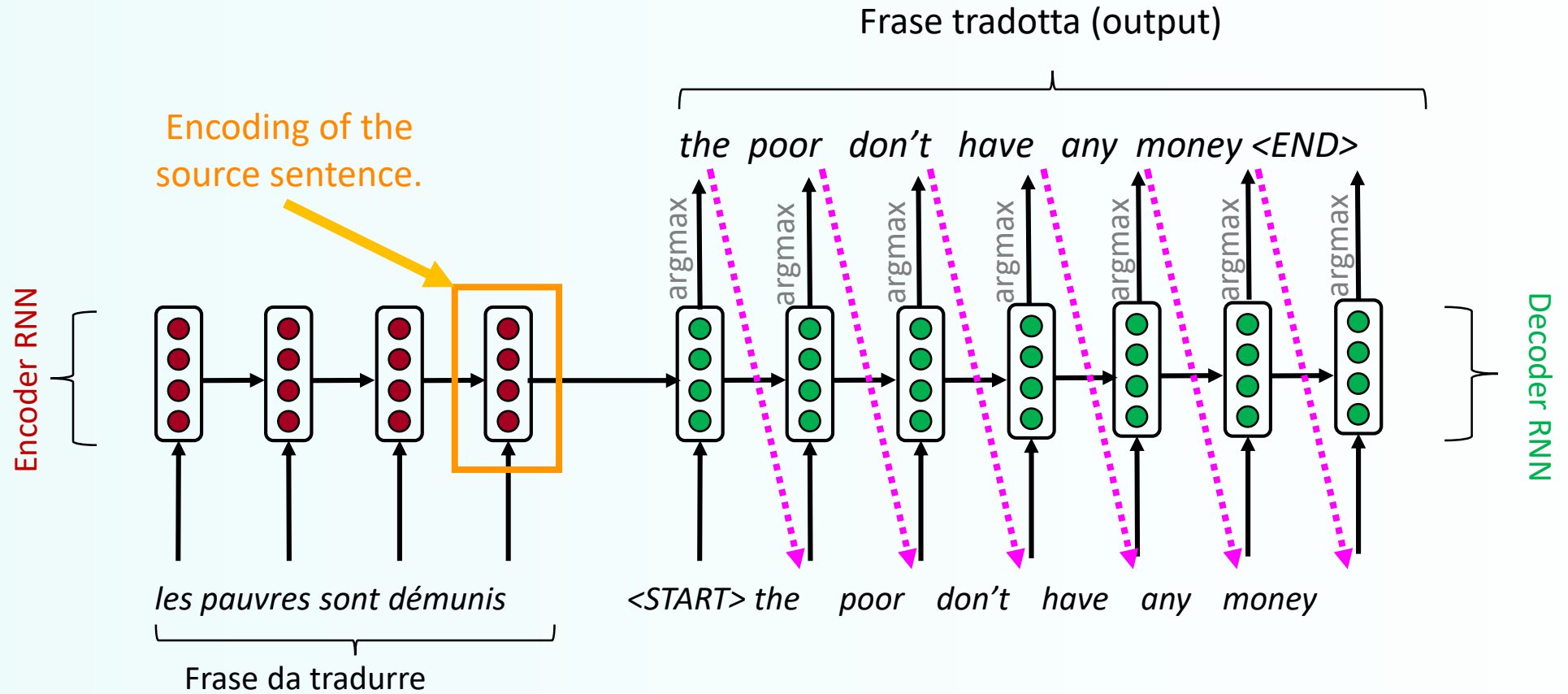
[Edinburgh En-De WMT newstest2013 Cased BLEU; NMT 2015 from U. Montréal]



MT è un problema risolto?

- **Per nulla!**
- Many difficulties remain:
 - **Out-of-vocabulary** words
 - **Domain mismatch** between train and test data
 - Maintaining **context** over longer text
 - **Low-resource** language pair

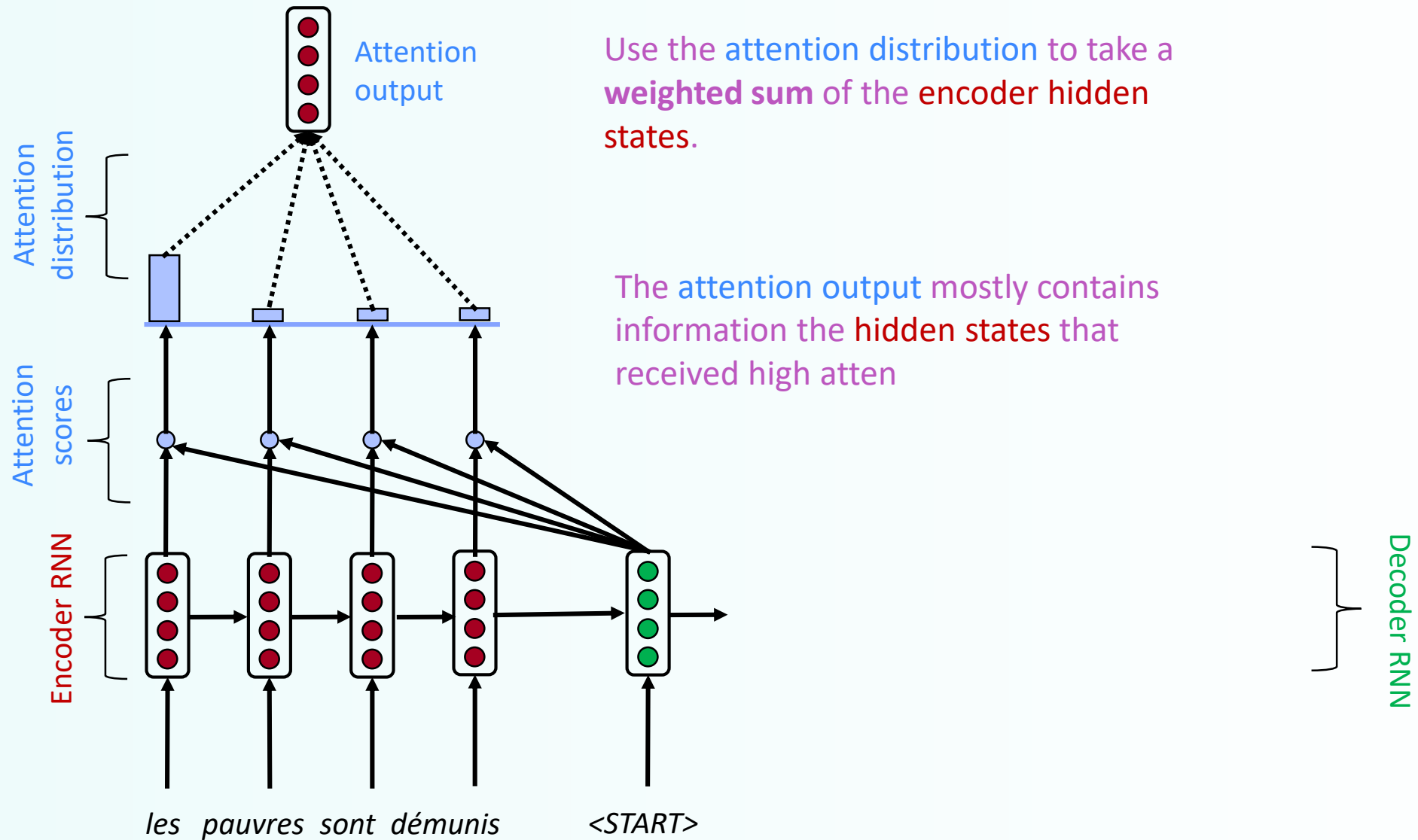
Limiti di questa Architettura?



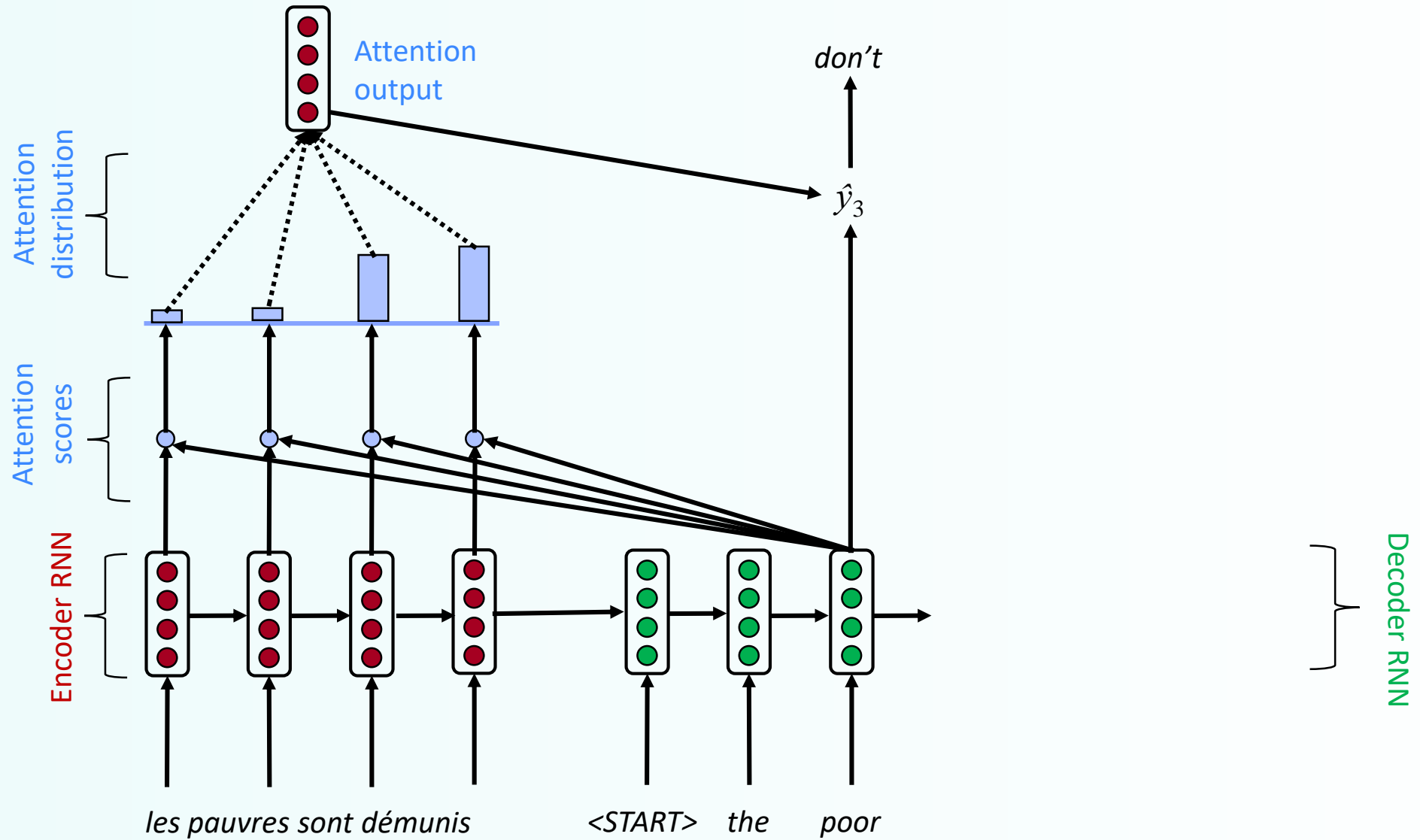
Meccanismo di Attention

- **Attention** provides a solution to the bottleneck problem.
- **Idea di base**: on each step of the decoder, *focus on a particular part* of the source sequence

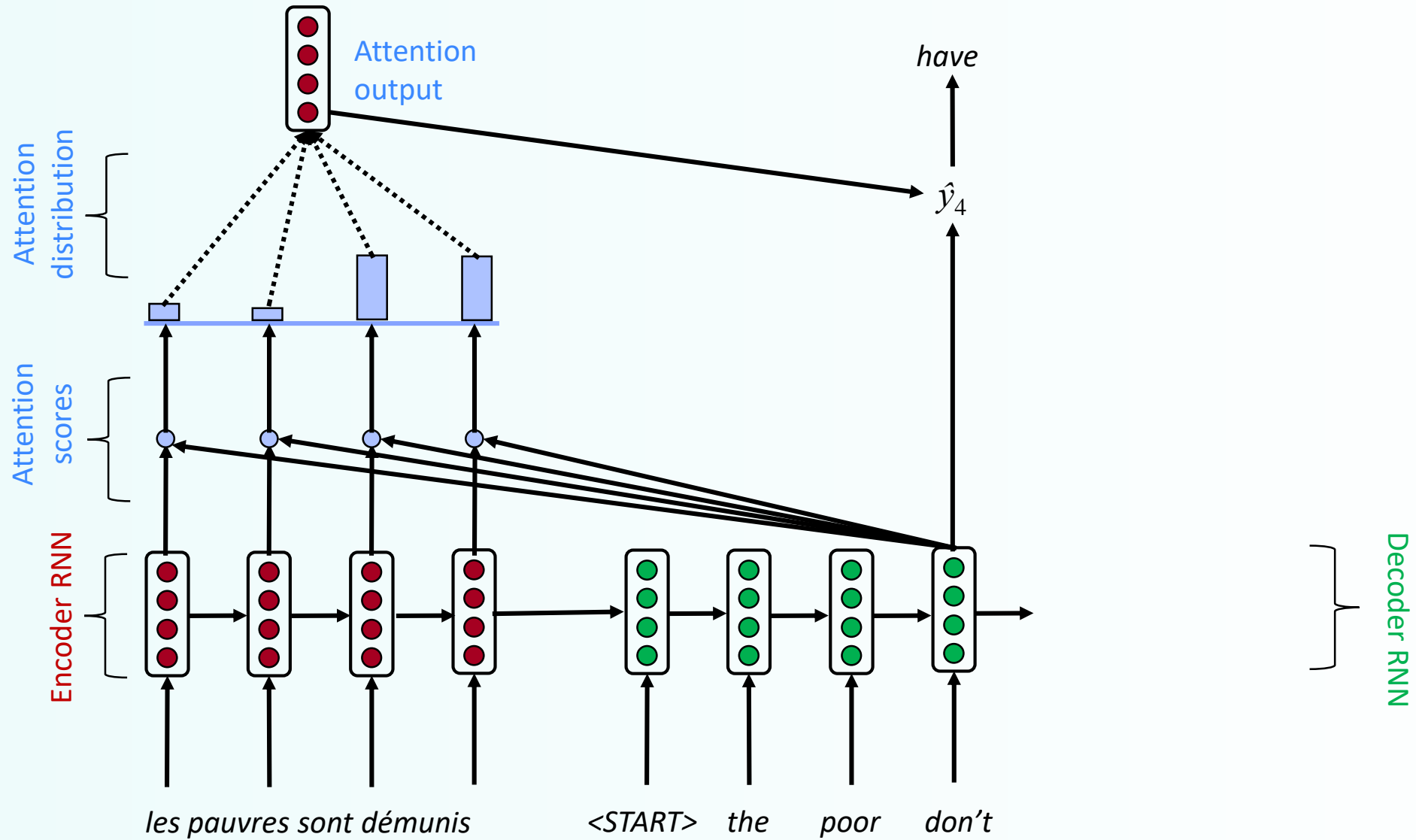
Sequence-to-sequence with attention



Sequence-to-sequence with attention



Sequence-to-sequence with attention



Attention è efficace

- Attention **migliora significativamente la NMT**
- Attention **risolve il problema del collo di bottiglia**
- Attention **reduce il vanishing gradient problem**
- Attention fornisce **interpretabilità**
 - Possiamo ricavare da quali parole della fonte ha origine la traduzione
 - Si ottiene **gratis l'allineamento tra parole!**
 - La rete lo impara da sè

A diagram illustrating attention weights between an English source sentence and a French target sentence. The English words are 'The', 'poor', 'don't', 'have', 'any', and 'money'. The French words are 'Les', 'pauvres', 'sont', and 'démunis'. A green arrow points to the first row of the matrix, which corresponds to the English word 'The'. The matrix shows that 'The' is aligned with 'Les', 'poor' with 'pauvres', 'don't' with 'sont', 'have' with 'démunis', 'any' with 'démunis', and 'money' with 'démunis'.

	Les	pauvres	sont	démunis
→ The	■			
poor		■		
don't			■	■
have			■	■
any			■	■
money			■	■

E il senso delle frasi?

- NT non tenta nemmeno di afferrare il senso delle frasi
- Capisce almeno la grammatica?
 - Nemmeno, anche se qualcuno ha provato a utilizzarla

Conclusioni

Abbiamo raccontato la storia della Machine Translation (MT)

- Dal 2014, Neural MT ha rapidamente sostituito Sistemi intricate di Statistical (Phrase Based) MT
- Lo stato dell'arte è dato da modelli Sequence-to-sequence con Attention (uses 2 RNNs)
- NMT fa sorprendentemente bene, considerate che ancora non capisce davvero i testi che traduce
- La tecnologia è solo ai primi passi e continuerà ad evolvere