

JST-NSF 国際連携シンポジウム 2019.12.19

言語生成技術 — データを言葉で記述する —

Natural Language Generation

— Describing Data with Words —



Hiroya Takamura

AIST/Tokyo Institute of Technology

さきがけ「社会デザイン」

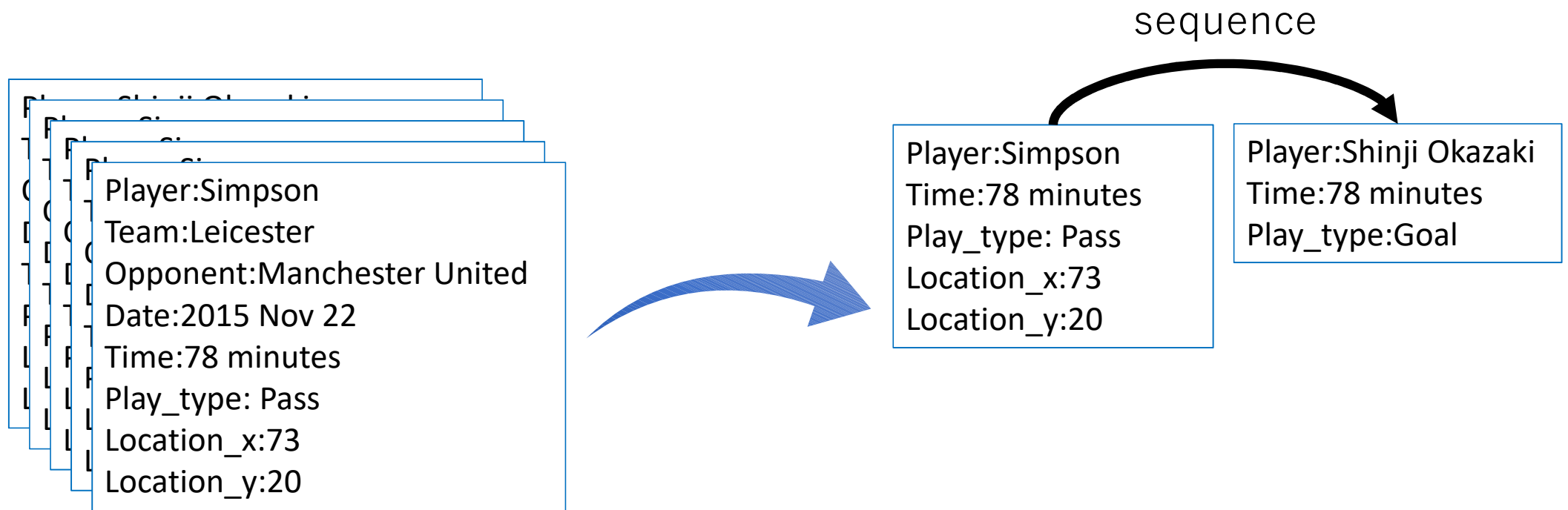
言語生成の伝統的な枠組み

Conventional Framework for Text Generation

1. 内容選択 (content selection) : 何を言うかを決定
determine what to say
2. マイクロプランニング (micro-planning) : どう言うかを決定
determine how to say
3. 表層化 (surface realization) : テキストを生成
generate text

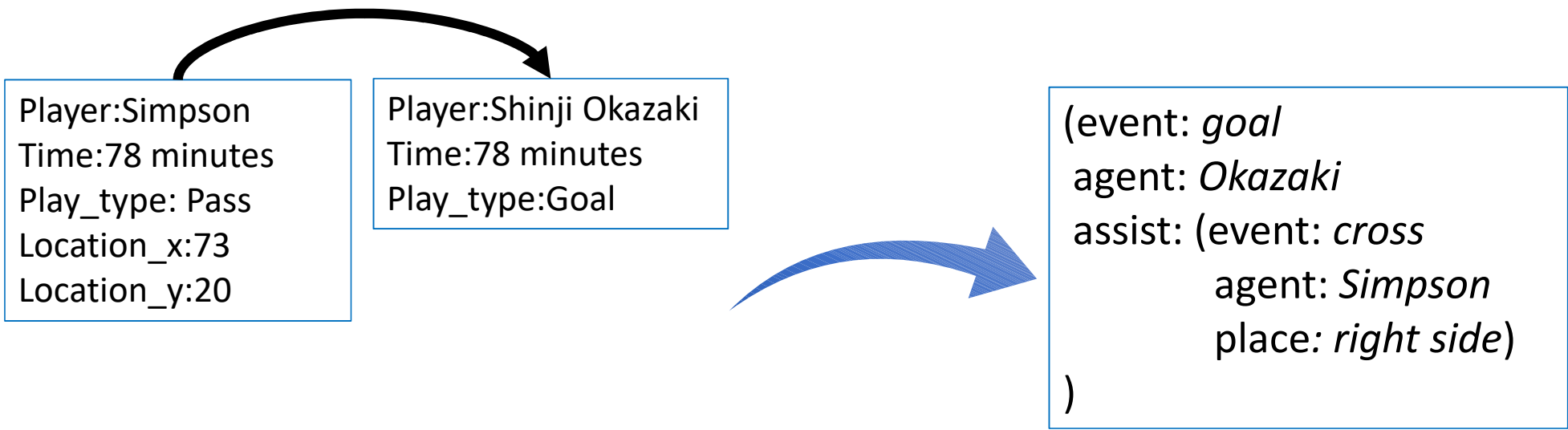
Conventional Framework for Text Generation

Content Selection : what to say



Conventional Framework for Text Generation

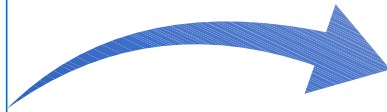
Micro-Planning : how to say sequence



Conventional Framework for Text Generation

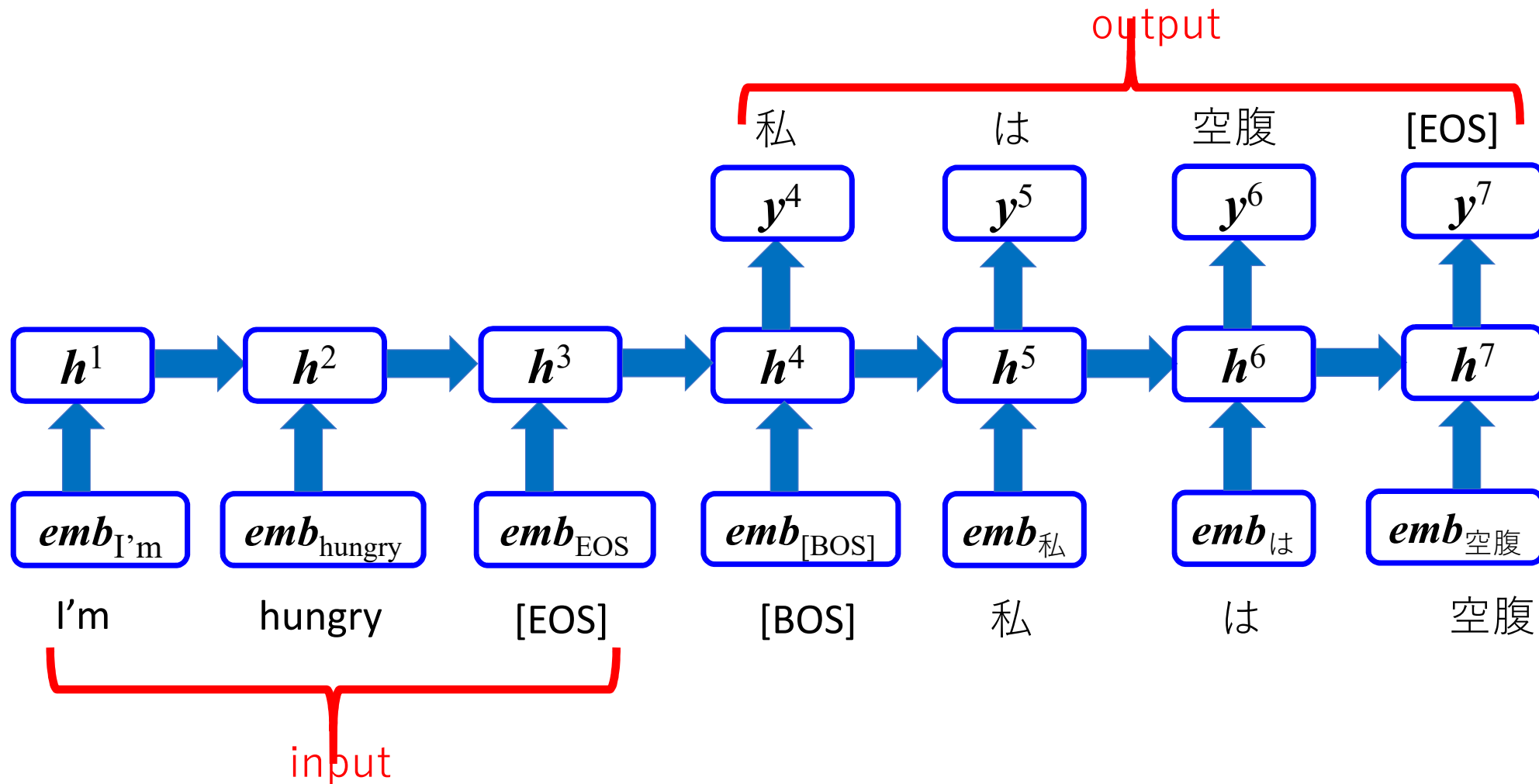
Surface Realization : generate text

```
(event: goal  
agent: Okazaki  
assist: (event: cross  
         agent: Simpson  
         place: right side)  
)
```

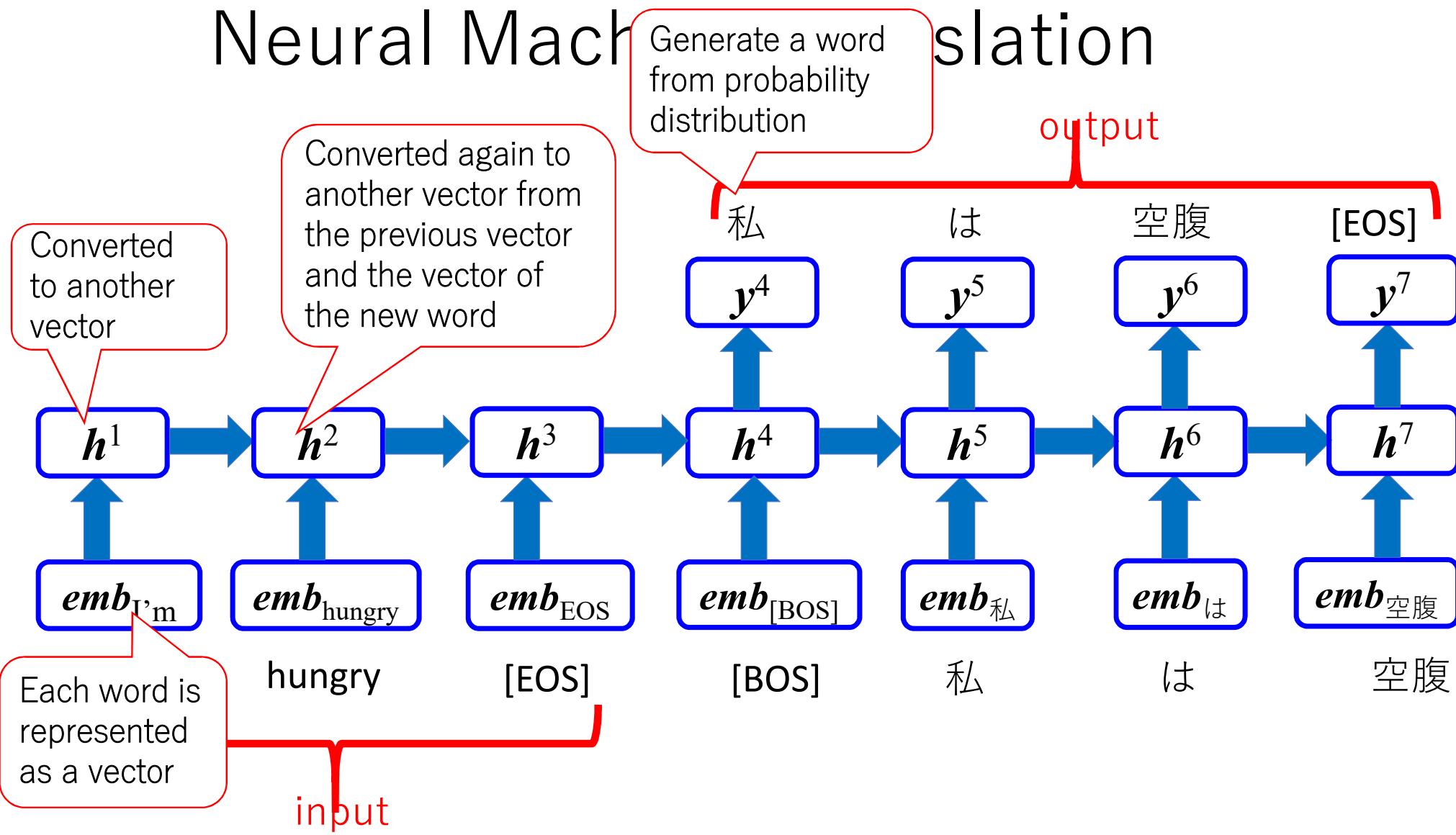


*Okazaki's goal assisted by
Simpson's cross from right side*

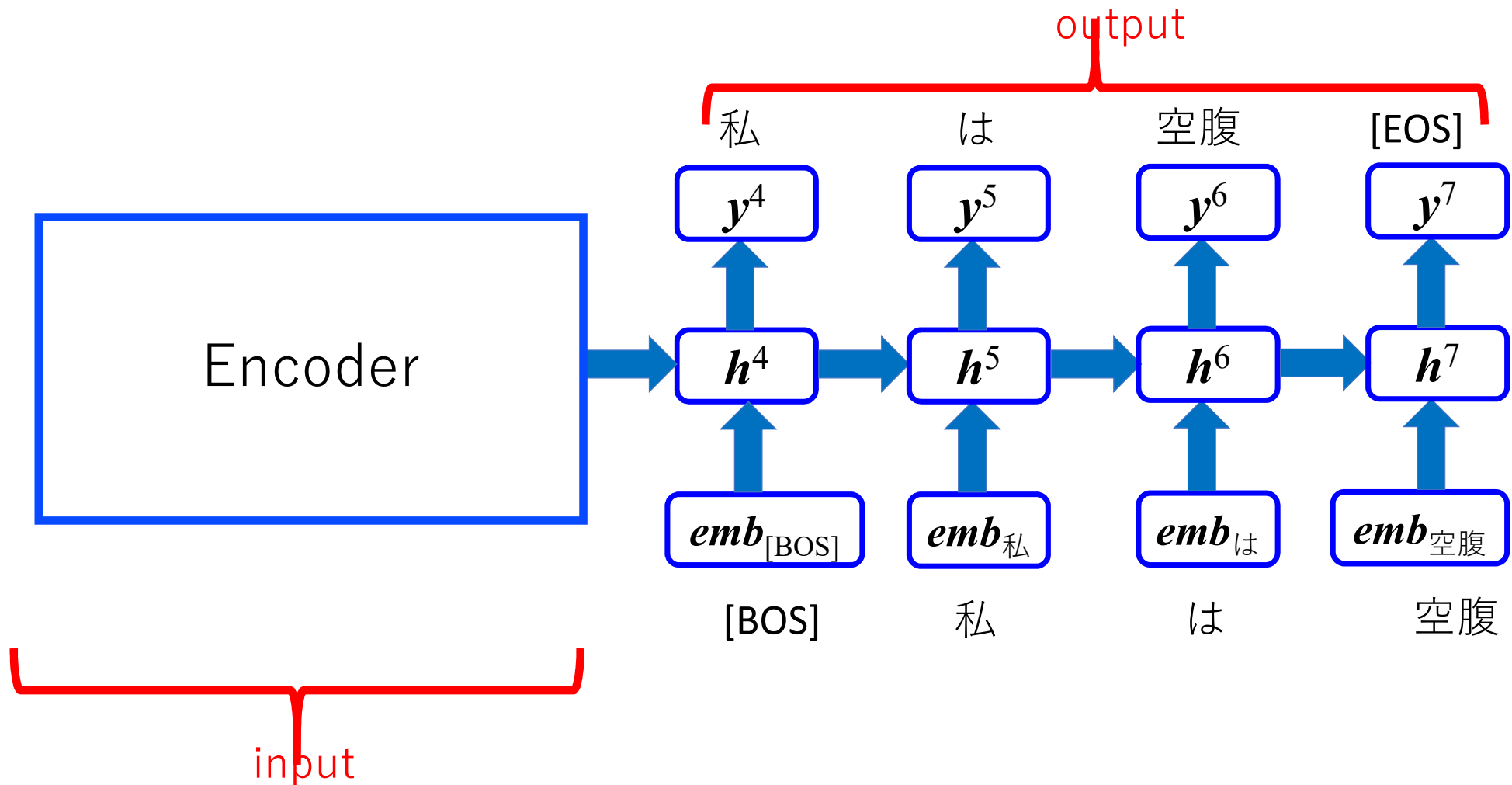
Neural Machine Translation



Neural Machine Translation



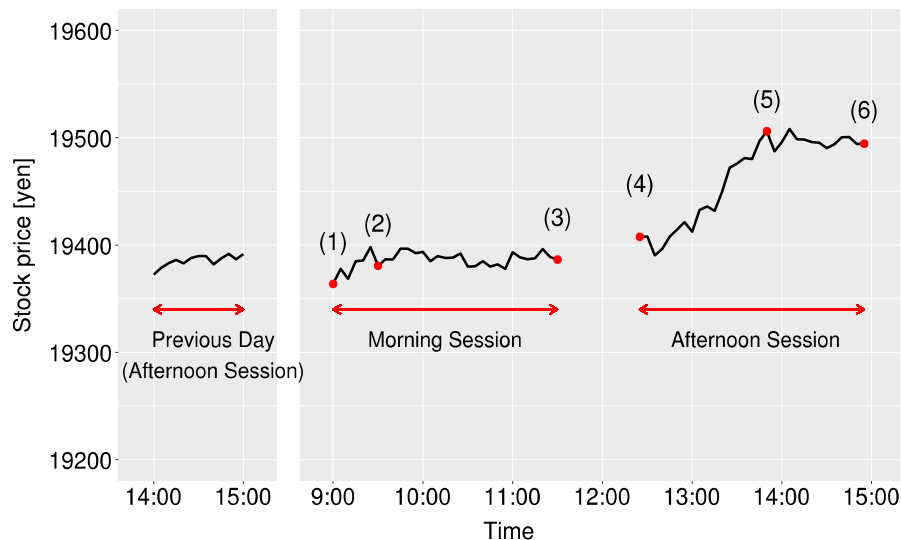
Encoder-Decoder



株価時系列データからの市況コメント生成

Generating market comments from stock prices

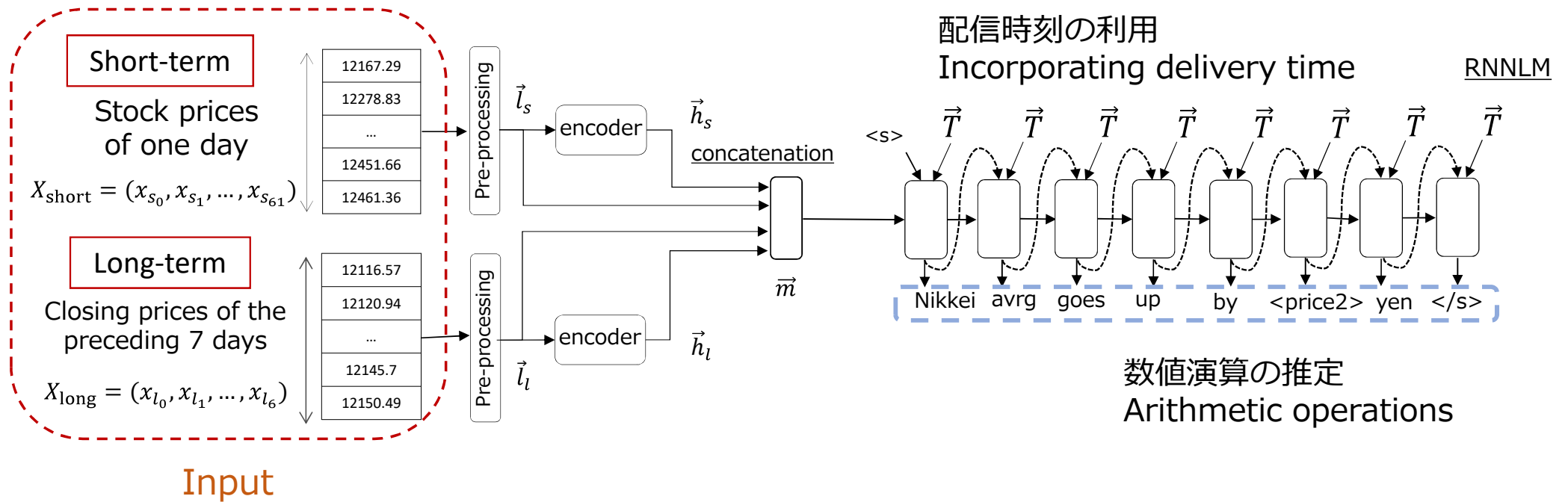
日経平均 (Nikkei stock average)



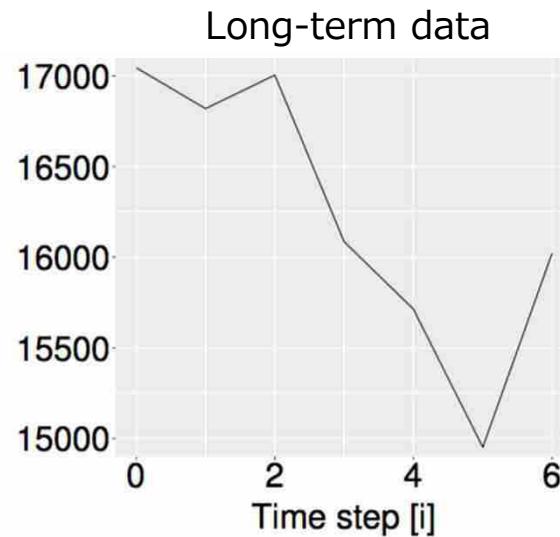
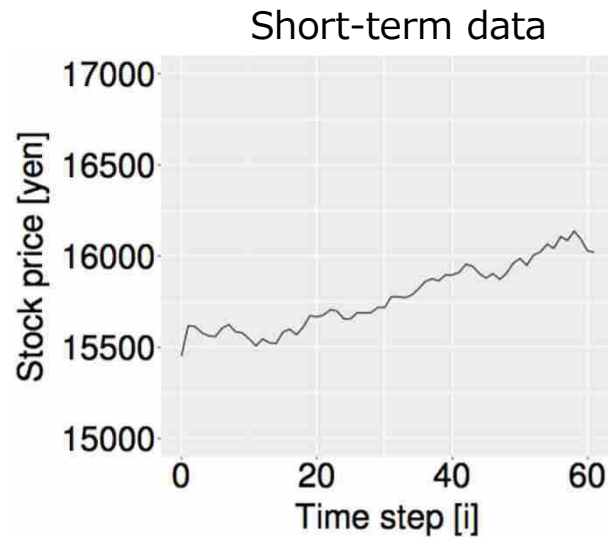
市況コメント (Market comments)

(1)	09:00	日経平均、続落で始まる Nikkei opens with a continual fall
(2)	09:29	日経平均、上げに転じる Nikkei turns to rise
(3)	11:30	日経平均、続落 前引けは5円安の19386円 Nikkei continues to fall. The closing price of the morning session decreases by 5 yen to 19,386 yen
(4)	12:30	日経平均、午後は上昇で始まる Nikkei rises at the beginning of the afternoon session
(5)	13:54	日経平均、上げ幅100円超える Nikkei gains more than 100 yen.
(6)	15:00	日経平均、反発 大引けは102円高の19494円 Nikkei rebounds and closes up 102 yen to 19,494 yen

市況コメント生成のためのニューラルネットワーク Neural Network for Generating Market Comments



出力例 Output Example



配信時刻 Posting time
2016/2/15 15:00

日経平均、大幅反発 大引けは1,069円高の16,022円

Nikkei significantly rebounds.

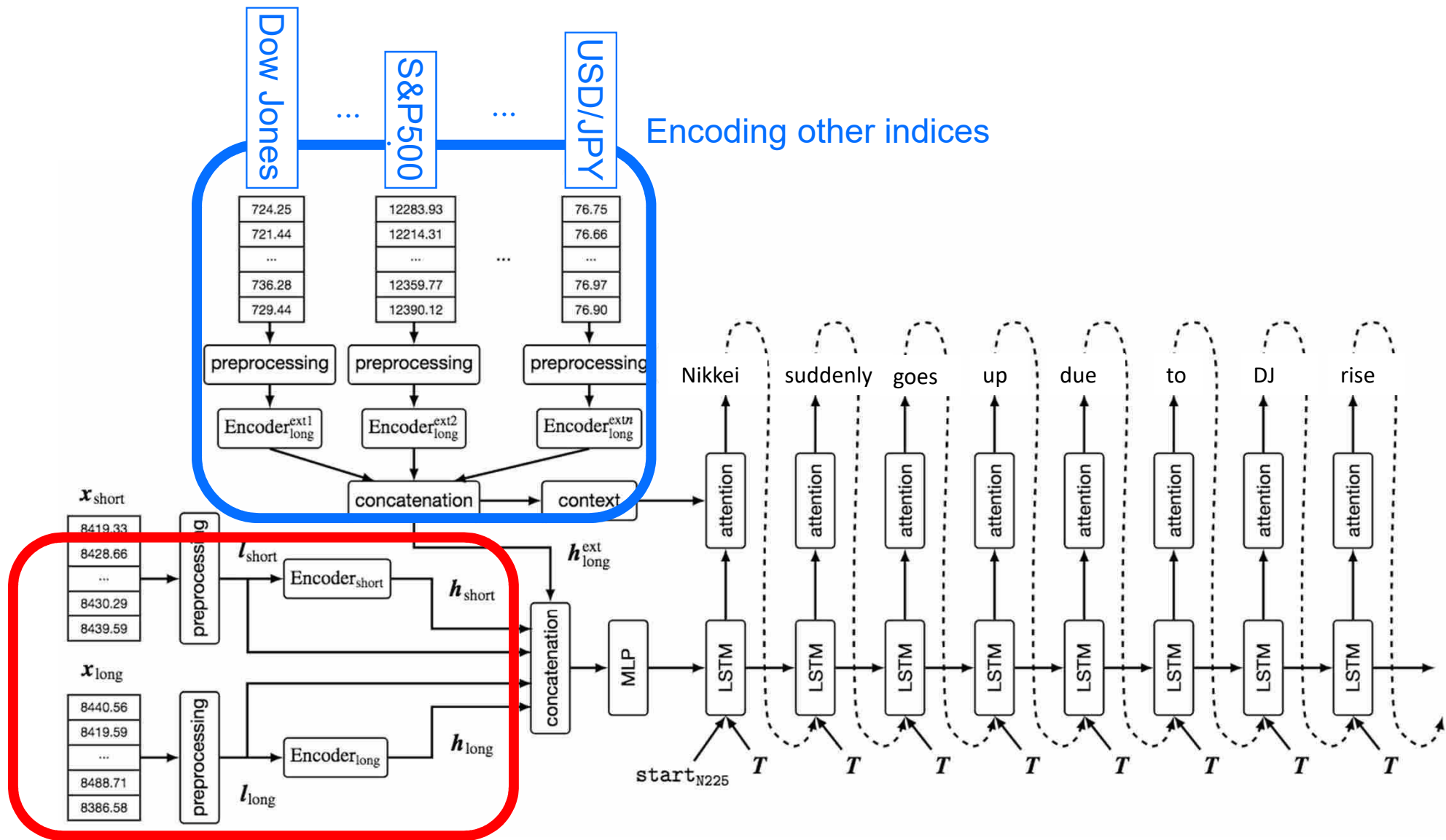
The closing price is 16,022 yen, which is 1,069 yen higher.

市況コメント生成における外部要因の記述

Generating Market Comments Referring to External Resources

日経平均だけでなく、Dow Jones平均、円ドル為替、アジア株平均などを見る必要がある

We need to look at other indices: Dow Jones, USD/JPY foreign exchange rate, Asian stock price, ...



Encoding the target index (Nikkei)

出力例 Output Example

人間が書いたテキスト text written by human:

日経平均、続落で始まる 欧米株安や円高進行で

Nikkei average starts with a continual fall, due to the decline of stock prices in US and Europe and the strong yen.

AIが書いたテキスト text written by AI:

日経平均、続落で始まる 米株安で、円高が重荷

Nikkei average starts with a continual fall, due to the decline of US stock prices and the strong yen.

※ 東工大との共同研究 (joint work with Tokyo Institute of Technology)

サッカーのプレイデータからのテキスト速報生成 Generating live soccer-match commentary from play data

```
<Q id="1724945974" qualifier_id="157"/>
<Q id="225210807" qualifier_id="56" value="Center"/>
<Q id="1096438506" qualifier_id="141" value="73.9"/>
</Event>
▼<Event id="2072246161" event_id="3" type_id="1" period_id
timestamp="2013-11-23T17:30:35.654" last_modified="2013-1
<Q id="1896741342" qualifier_id="3"/>
<Q id="896027247" qualifier_id="140" value="46.8"/>
<Q id="357823138" qualifier_id="212" value="24.2"/>
<Q id="2067037058" qualifier_id="56" value="Back"/>
<Q id="1083716141" qualifier_id="141" value="35.7"/>
<Q id="2077976868" qualifier_id="213" value="0.3"/>
</Event>
<Event id="767235142" event_id="4" type_id="49" period_id
timestamp="2013-11-23T17:30:37.588" last_modified="2013-1
▼<Event id="1745340463" event_id="5" type_id="1" period_id
timestamp="2013-11-23T17:30:37.637" last_modified="2013-1
<Q id="858656129" qualifier_id="213" value="2.2"/>
<Q id="1850023696" qualifier_id="140" value="37.0"/>
<Q id="912347679" qualifier_id="141" value="58.4"/>
<Q id="507101334" qualifier_id="56" value="Back"/>
<Q id="1014878810" qualifier_id="212" value="18.5"/>
</Event>
▼<Event id="1353554706" event_id="6" type_id="1" period_id
timestamp="2013-11-23T17:30:38.476" last_modified="2013-1
<Q id="1463575318" qualifier_id="140" value="44.0"/>
<Q id="196597581" qualifier_id="141" value="78.4"/>
<Q id="2102939476" qualifier_id="213" value="1.1"/>
<Q id="880472161" qualifier_id="56" value="Back"/>
<Q id="1430823066" qualifier_id="212" value="15.4"/>
</Event>
▼<Event id="784379380" event_id="7" type_id="1" period_id
timestamp="2013-11-23T17:30:43.402" last_modified="2013-1
```



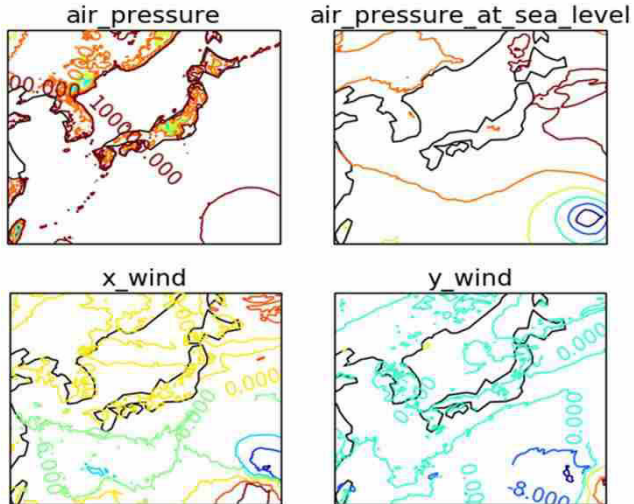
Luis Suarez, on the right wing, tries to lay the ball off to captain Steven Gerrard, but the striker's pass doesn't quite have enough accuracy on it.

※ 東工大における成果 (research at Tokyo Institute of Technology)

Generating Weather Reports from Simulated Meteorological Data

Meteorological data

(air pressure, wind, cloud amount, humidity, precipitation, etc)



Weather report

This week starts with cold rain or snow.

Simulated time-series data

ボックススコアからの試合サマリの生成

Generating sports-match summary from box score

Box score

Summary

	CITY	WINS	LOSSES	PTS	QTR1	QTR2	QTR3	QTR4	AST	REB	TOV	FG_PCT	FT_PCT	FG3_PCT
Hawks	Atlanta	7	15	95	25	21	17	32	20	33	16	43	80	41
Heat	Miami	11	12	103	29	24	25	25	27	47	14	49	87	21

	H/V	POS	MIN	PTS	REB	AST	BLK	TO	PF	STL	DREB	OREB	FGM	FGA	FG_PCT	FTM	FTA	FT_PCT	FG3M	FG3A	FG3_PCT
PLAYER_NAME																					
Hassan Whiteside	H	C	36	8	12	0	1	4	5	0	9	3	4	12	33	0	0	0	0	0	0
Wayne Ellington	H	F	41	19	3	2	0	0	0	0	3	0	7	15	47	1	1	100	4	9	44
Josh McRoberts	H	F	30	2	2	4	0	5	2	1	1	1	1	5	20	0	0	0	0	1	0
Rodney McGruder	H	G	31	11	5	5	1	1	1	2	2	3	3	8	38	3	5	60	2	5	40
Goran Dragic	H	G	35	21	2	4	1	3	4	1	2	0	8	17	47	4	4	100	1	4	25
James Johnson	H																				
Josh Richardson	H																				
Justise Winslow	H																				
Tyler Johnson	H		35	27	2	5	1	0	1	1	2	0	8	16	50	7	8	88	4	5	80
Willie Reed	H		12	4	2	0	0	0	1	1	1	1	2	2	100	0	0	0	0	0	0
Derrick Williams	H		20	3	5	0	2	3	2	0	5	0	1	5	20	1	2	50	0	3	0
Luke Babbitt	H																				
Udonis Haslem	H																				
Dwight Howard	V	C	35	23	17	4	2	1	2	1	10	7	9	11	82	5	6	83	0	0	0
Thabo Sefolosha	V	F	29	10	5	5	0	0	0	1	2	3	5	11	46	0	0	0	0	3	0
Paul Millsap	V	F	32	21	9	2	3	2	2	4	8	1	8	12	67	5	6	83	0	2	0
Tim Hardaway Jr.	V	G	22	7	0	2	0	2	3	1	0	0	3	9	33	0	0	0	1	3	33
Dennis Schroeder	V	G	31	17	4	7	1	3	2	1	3	1	8	15	53	0	0	0	1	2	50
Kyle Korver	V		34	9	3	5	0	1	3	0	3	0	3	9	33	1	1	100	2	6	33
Mike Muscala	V		24	8	3	0	2	2	1	0	3	0	4	8	50	0	0	0	0	1	0
Kris Humphries	V		4	0	2	0	0	1	1	0	0	2	0	2	0	0	0	0	0	0	0
Malcolm Delaney	V		17	4	2	1	0	2	3	1	2	0	2	6	33	0	0	0	0	1	0
Taurean Prince	V		12	4	2	1	1	0	3	0	2	0	1	5	20	2	2	100	0	1	0

The Atlanta Hawks defeated the Miami Heat , 103 - 95 , at Philips Arena on Wednesday . Atlanta was in desperate need of a win and they were able to take care of a shorthanded Miami team here . Defense was key for the Hawks , as they held the Heat to 42 percent shooting and forced them to commit 16 turnovers . Atlanta also dominated in the paint , winning the rebounding battle , 47 - 34 , and outscoring them in the paint 58 - 26. The Hawks shot 49 percent from the field and assisted on 27 of their 43 made baskets . This was a near wire - to - wire win for the Hawks , as Miami held just one lead in the first five minutes . Miami (7 - 15) are as beat - up as anyone right now and it 's taking a toll on the heavily used starters . Hassan Whiteside really struggled in this game , as he amassed eight points , 12 rebounds and one blocks on 4 - of - 12 shooting . Tyler Johnson was the team 's best player , as he tallied 27 points and five assists off the bench . Goran Dragic continued his good form , scoring 21 points and handing out four assists . Rodney McGruder got the start for Justise Winslow and finished with 11 points , five rebounds and five assists . It does n't get any easier for the Heat , as they play the Cleveland Cavaliers on the road for their next game . Atlanta (11 - 12) had lost seven - straight headed into this game , so to say this win was needed would be a drastic understatement . Dwight Howard took a lot of the load on his broad shoulders , as he tallied 23 points , 17 rebounds and four assists . It was n't just offense though , as he held the dominant Hassan Whiteside in check . Paul Millsap also played great , accumulating 21 points , nine rebounds , four steals and three blocks in the victory . Dennis Schroeder was the only other player in double figures , accruing 17 points , four rebounds and seven assists . Atlanta will look to build off

通常のencoder-decoderは長い文書が苦手

The naïve encoder-decoder is not good at long documents

The Utah Jazz (38 - 26) defeated the Houston Rockets (38 - 26) 117 - 91 on Wednesday at Energy Solutions Arena in Salt Lake City . The Jazz got out to a quick start in this one , out - scoring the Rockets 31 - 15 in the first quarter alone . Along with the quick start , the Rockets were the superior shooters in this game , going 54 percent from the field and 43 percent from the three - point line , while the Jazz went 38 percent from the floor and a meager 19 percent from deep . The Rockets were able to out - rebound the Rockets 49 - 49 , giving them just enough of an advantage to secure the victory in front of their home crowd . The Jazz were led by the duo of Derrick Favors and James Harden . Favors went 2 - for - 6 from the field and 0 - for - 1 from the three - point line to score a game - high of 15 points , while also adding four rebounds and four assists

blue : correct entities and numbers
red : wrong ones

かなり誤りが多い
数字のエラーや、意味を成さない文など

Lots of errors.
Errors in numbers.
Some sentences do not make sense.

言語生成の伝統的な枠組み

Conventional Framework for Text Generation

1. 内容選択 (content selection) : 何を言うかを決定
determine what to say
2. マイクロプランニング (micro-planning) : どう言うかを決定
determine how to say
3. 表層化 (surface realization) : テキストを生成
generate text

言語生成の伝統的な枠組み Conventional Framework for T

この考え方を援用すれば、
ニューラルネットワークで長い文書を生成できるのでは？
Simulated time-series data

1. 内容選択 (content selection) : 何を言うかを決定
determine what to say
2. マイクロプランニング (micro-planning) : どう言うかを決定
determine how to say
3. 表層化 (surface realization) : テキストを生成
generate text

ボックススコアからの試合サマリの生成

Generating sports-match summary from box score

言及すべき情報を
明示的に選択する

Box score

Summary

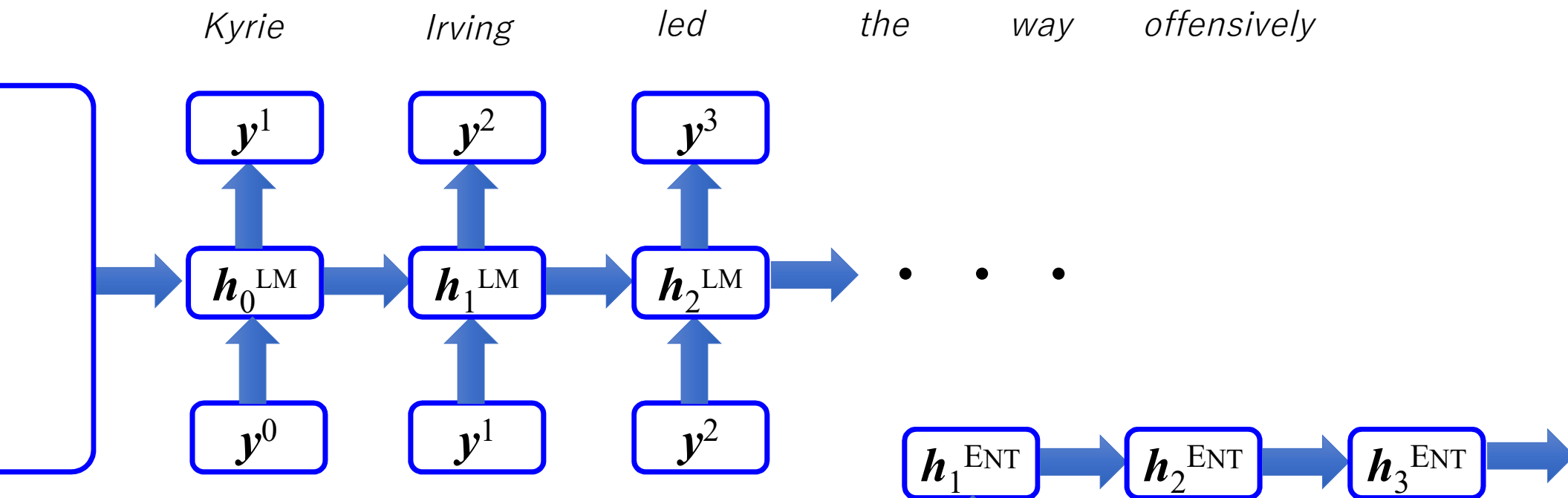
	CITY	WINS	LOSSES	PTS	QTR1	QTR2	QTR3	QTR4	AST	REB	TOV	FG_PCT	FT_PCT	FG3_PCT
Hawks	Atlanta	7	15	95	25	21	17	32	20	33	16	43	80	41
Heat	Miami	11	12	107	29	24	25	25	27	47	14	49	87	21

	H/V	POS	MIN	PTS	REB	AST	BLK	TO	PF	STL	DREB	OREB	FGM	FGA	FG_PCT	FTM	FTA	FT_PCT	FG3M	FG3A	FG3_PCT
PLAYER_NAME																					
Hassan Whiteside	H	C	36	8	12	0	1	4	5	0	9	3	4	12	33	0	0	0	0	0	0
Wayne Ellington	H	F	41	19	3	2	0	0	0	0	3	0	7	15	47	1	1	100	4	9	44
Josh McRoberts	H	F	30	2	2	4	0	5	2	1	1	1	1	5	20	0	0	0	0	1	0
Rodney McGruder	H	G	31	11	5	5	1	1	1	2	2	3	3	8	38	3	5	60	2	5	40
Goran Dragic	H	G	35	21	2	4	1	3	4	1	2	0	8	17	47	4	4	100	1	4	25
James Johnson	H																				
Josh Richardson	H																				
Justise Winslow	H																				
Tyler Johnson	H		35	27	2	5	1	0	1	1	2	0	8	16	50	7	8	88	4	5	80
Willie Reed	H		12	4	2	0	0	0	1	1	1	1	2	2	100	0	0	0	0	0	0
Derrick Williams	H		20	3	5	0	2	3	2	0	5	0	1	5	20	1	2	50	0	3	0
Luke Babbitt	H																				
Udonis Haslem	H																				
Dwight Howard	V	C	35	23	17	4	2	1	2	1	10	7	9	11	82	5	6	83	0	0	0
Thabo Sefolosha	V	F	29	10	5	5	0	0	0	1	2	3	5	11	46	0	0	0	0	3	0
Paul Millsap	V	F	32	21	9	2	3	2	2	4	8	1	8	12	67	5	6	83	0	2	0
Tim Hardaway Jr.	V	G	22	7	0	2	0	2	3	1	0	0	3	9	33	0	0	0	1	3	33
Dennis Schroeder	V	G	31	17	4	7	1	3	2	1	3	1	8	15	53	0	0	0	1	2	50
Kyle Korver	V		34	9	3	5	0	1	3	0	3	0	3	9	33	1	1	100	2	6	33
Mike Muscala	V		24	8	3	0	2	2	1	0	3	0	4	8	50	0	0	0	0	1	0
Kris Humphries	V		4	0	2	0	0	1	1	0	0	2	0	2	0	0	0	0	0	0	0
Malcolm Delaney	V		17	4	2	1	0	2	3	1	2	0	2	6	33	0	0	0	0	1	0
Taurean Prince	V		12	4	2	1	1	0	3	0	2	0	1	5	20	2	2	100	0	1	0

The Atlanta Hawks defeated the Miami Heat , 103 - 95 , at Philips Arena on Wednesday . Atlanta was in desperate need of a win and they were able to take care of a shorthanded Miami team here . Defense was key for the Hawks , as they held the Heat to 42 percent shooting and forced them to commit 16 turnovers . Atlanta also dominated in the paint , winning the rebounding battle , 47 - 34 , and outscoring them in the paint 58 - 26. The Hawks shot 49 percent from the field and assisted on 27 of their 43 made baskets . This was a near wire - to - wire win for the Hawks , as Miami held just one lead in the first five minutes . Miami (7 - 15) are as beat - up as anyone right now and it 's taking a toll on the heavily used starters . Hassan Whiteside really struggled in this game , as he amassed eight points , 12 rebounds and one blocks on 4 - of - 12 shooting . Tyler Johnson was the team 's best player , as he tallied 27 points and five assists off the bench . Goran Dragic continued his good form , scoring 21 points and handing out four assists . Rodney McGruder got the start for Justise Winslow and finished with 11 points , five rebounds and five assists . It does n't get any easier for the Heat , as they play the Cleveland Cavaliers on the road for their next game . Atlanta (11 - 12) had lost seven - straight headed into this game , so to say this win was needed would be a drastic understatement . Dwight Howard took a lot of the load on his broad shoulders , as he tallied 23 points , 17 rebounds and four assists . It was n't just offense though , as he held the dominant Hassan Whiteside in check . Paul Millsap also played great , accumulating 21 points , nine rebounds , four steals and three blocks in the victory . Dennis Schroeder was the only other player in double figures , accruing 17 points , four rebounds and seven assists . Atlanta will look to build off

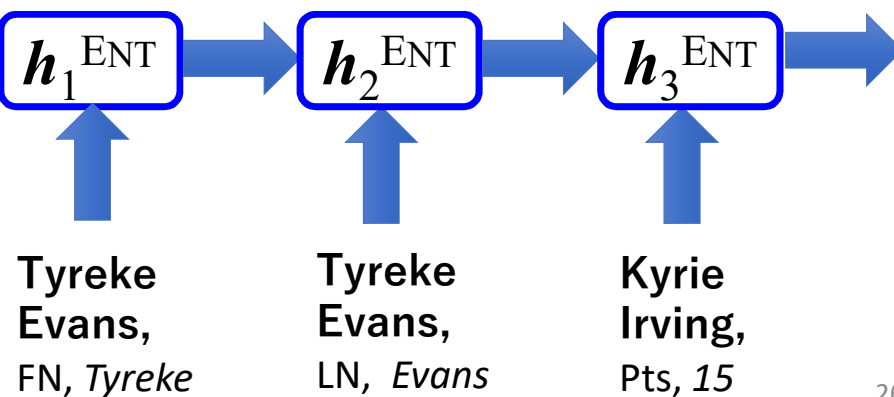
エンコーダ-デコーダ + エンティティトラッキング

言及したエンティティとその属性値の情報を記憶



例えば、次のような効果が期待できる

- Tyreke Evansの名前の次はTyreke Evansの得点数に言及する
- Tyreke Evansばかりについて何度も言及しない



The **Milwaukee Bucks** defeated the **New York Knicks**, **105-104**, at Madison Square Garden on Wednesday. The **Knicks (16-19)** checked in to Wednesday's contest looking to snap a five-game losing streak and heading into the fourth quarter, they looked like they were well on their way to that goal. . . . **Antetokounmpo** led the Bucks with **27** points, **13** rebounds, **four** assists, a steal and **three** blocks, his second consecutive double-double. **Greg Monroe** actually checked in as the second-leading scorer and did so in his customary bench role, posting **18** points, along with **nine** boards, **four** assists, **three** steals and a block. **Jabari Parker** contributed **15** points, **four** rebounds, **three** assists and a steal. **Malcolm Brogdon** went for **12** points, **eight** assists and **six** rebounds. **Mirza Teletovic** was productive in a reserve role as well, generating **13** points and a rebound. . . . The Bucks and Knicks face off once again in the second game of the home-and-home series, with the meeting taking place Friday night in Milwaukee.

Summary written by human writer

- **Green** : correct numbers
- **Orange** : wrong numbers

The **Milwaukee Bucks** defeated the **New York Knicks**, **105-104**, at Madison Square Garden on Saturday. The **Bucks (18-16)** checked in to Saturday's contest with a well, outscoring the **Knicks (16-(19))** by a margin of 39-19 in the first quarter. However, New York by just a 25-foot lead at the end of the first quarter, the Bucks were able to pull away, as they outscored the Knicks by a 59-46 margin into the second. 45 points in the third quarter to seal the win for New York with the rest of the starters to seal the win. The Knicks were led by **Giannis Antetokounmpo**, who tallied a game-high **27** points, to go along with **13** rebounds, **four** assists, **three** blocks and a steal. The game was a crucial night for the Bucks' starting five, as the duo was the most effective shooters, as they posted Milwaukee to go on a pair of low low-wise (Carmelo Anthony) and Malcolm Brogdon. **Anthony** added **11** rebounds, **seven** assists and **two** steals to his team-high scoring total. **Jabari Parker** was right behind him with **15** points, **four** rebounds, **three** assists and a block. **Greg Monroe** was next with a bench-leading **18** points, along with **nine** rebounds, **four** assists and **three** steals. **Brogdon** posted **12** points, **eight** assists, **six** rebounds and a steal. **Derrick Rose** and **Courtney Lee** were next with a pair of **{11 / 11}** -point efforts. **Rose** also supplied **four** assists and **three** rebounds, while **Lee** complemented his scoring with **three** assists, a rebound and a steal. **John Henson** and **Mirza Teletovic** were next with a pair of **{two / two}** -point efforts. **Teletovic** also registered **13** points, and he added a rebound and an assist. **Jason Terry** supplied **eight** points, **three** rebounds and a pair of steals. The Cavs remain in last place in the Eastern Conference's Atlantic Division. They now head home to face the Toronto Raptors on Saturday night.

Summary written by our AI

The **Milwaukee Bucks** defeated the **New York Knicks**, **105-104**, at Madison Square Garden on Wednesday evening. The **Bucks (18-16)** have been one of the hottest teams in the league, having won five of their last six games, and they have now won six of their last eight games. The **Knicks (16-19)** have now won six of their last six games, as they continue to battle for the eighth and final playoff spot in the Eastern Conference. **Giannis Antetokounmpo** led the way for Milwaukee, as he tallied **27** points, **13** rebounds, **four** assists, **three** blocked shots and **one** steal, in **39** minutes. **Jabari Parker** added **15** points, **four** rebounds, **three** assists, **one** steal and **one** block, and **6-of-8** from long range. **John Henson** added **two** points, **two** rebounds, **one** assist, **three** steals and **one** block. **John Henson** was the only other player to score in double digits for the Knicks, with **15** points, **four** assists, **three** rebounds and **one** steal, in **33** minutes. The Bucks were led by **Derrick Rose**, who tallied **15** points, **four** assists, **three** rebounds and **one** steal in **33** minutes. **Willy Hernangomez** started in place of Porzingis and finished with **15** points, **four** assists, **three** rebounds and **one** steal in **33** minutes. **Willy Hernangomez** started in place of Jose Calderon (knee) and responded with **one** rebound and **one** block. The Knicks were led by their starting backcourt of **Carmelo Anthony** and **Carmelo Anthony**, but combined for just **13** points on 5-of-16 shooting. The Bucks next head to Philadelphia to take on the Sixers on Friday night, while the Knicks remain home to face the Los Angeles Clippers on Wednesday.

Summary written by Puduppully (2019)'s AI

The **Milwaukee Bucks** defeated the **New York Knicks**, **105-104**, at Madison Square Garden on Saturday. The **Bucks (18-16)** checked in to Saturday's contest with a well, outscoring the **Knicks (16-19)** by a margin of 39-19 in the first quarter. However, New York by just a 25-foot lead at the end of the first quarter, the Bucks were able to pull away, as they outscored the Knicks by a 59-46 margin into the second. 45 points in the third quarter to seal the win for New York with the rest of the starters to seal the win. The Knicks were led by **Giannis Antetokounmpo**, who tallied a game-high **27** points, to go along with **13** rebounds, **four** assists, **three** blocks and a steal. The game was a crucial night for the Bucks' starting five, as the duo was the most effective shooters, as they posted Milwaukee to go on a pair of low low-wise (Carmelo Anthony) and Malcolm Brogdon. **Anthony** added **11** rebounds, **seven** assists and **two** steals to his team-high scoring total. **Jabari Parker** was right behind him with **15** points, **four** rebounds, **three** assists and a block. **Greg Monroe** was next with a bench-leading **18** points, along with **nine** rebounds, **four** assists and **three** steals. **Brogdon** posted **12** points, **eight** assists, **six** rebounds and a steal. **Derrick Rose** and **Courtney Lee** were next with a pair of **{11 / 11}** -point efforts. **Rose** also supplied **four** assists and **three** rebounds, while **Lee** complemented his scoring with **three** assists, a rebound and a steal. **John Henson** and **Mirza Teletovic** were next with a pair of **{two / two}** -point efforts. **Teletovic** also registered **13** points, and he added a rebound and an assist. **Jason Terry** supplied **eight** points, **three** rebounds and a pair of steals. The Cavs remain in last place in the Eastern Conference's Atlantic Division. They now head home to face the Toronto Raptors on Saturday night.

Summary written by our AI