

文の生成

長尾真編、岩波講座ソフトウェア科学、
自然言語処理、9章、「文の生成」から

簡単な文の生成法

- ▶ 文のランダム生成
- ▶ スロット法
- ▶ 言い換え
- ▶ 文の変形
- ▶ 対話における文生成の例: ELIZA
- ▶ 格構造からの文の生成



文のランダム生成

- ▶ 文脈自由文法で、初期記号から出発して書き換え規則をランダムに適用する。(V.H.A.Yngve)
- ▶ 意味的な制約が考慮されていない。
- ▶ 意味要素を取り入れた文の生成(長尾1965)



文脈自由文法の例

$$G_k = (V_k, \Sigma_k, P_k, \text{文})$$

$$\Sigma_k = \{\text{きた, 文化, 伝わった, から, が}\}$$

$$V_k = \{\text{文, 後置詞句, 動詞, 名詞, 助詞}\}$$

$$\cup \Sigma_k$$

$$P_k = \{ \text{文} \rightarrow \text{後置詞句 文},$$

$$\text{文} \rightarrow \text{後置詞句 動詞},$$

$$\text{後置詞句} \rightarrow \text{名詞 助詞},$$

$$\text{後置詞句} \rightarrow \text{動詞 助詞},$$

$$\text{名詞} \rightarrow \text{きた}, \text{名詞} \rightarrow \text{文化}$$

$$\text{動詞} \rightarrow \text{きた}, \text{動詞} \rightarrow \text{伝わった}$$

$$\text{助詞} \rightarrow \text{から}, \text{助詞} \rightarrow \text{が} \}$$



スロット法 (例)

- (1) ただ今、<s1>線に到着しました列車は、<s2>時<s3>分発の<s4>行きの列車です。
- (2) <s2>時<s3>分発<s4>行き列車は<s1>線から発車します。
- (3) <s6><s7>号の停車駅は<s4>, ..., <s4>です。

<s1> : n番

<s4> : 駅名

<s2> : $n (0 \leq n < 23)$

<s5> : 特急、急行、普通

<s3> : $n (0 \leq n < 59)$

<s6> : ひかり、こだま

<s7> : 3桁までの数字

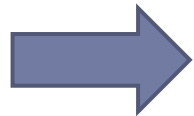
(1) <l> want to go to New York.

(2) <l> want to have a cup of <tea>.

l → Heのとき、want → wants

言い換え

▶ これは机です。



これは椅子です。

あれは机です。

それは椅子です。

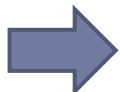
勉強＝勉学＝学習

庭＝庭園

父＝父親＝おやじ

私＝わたくし＝おれ＝ぼく＝わたし

彼は熱心に勉強している。



彼は熱心に学習している。

父はいつもうるさく注意する。



おやじはいつもうるさく注意する。

文の変形

変形文法(チョムスキー)

agent verb object  object be+verb-ed by agent

agent  object を verb する。

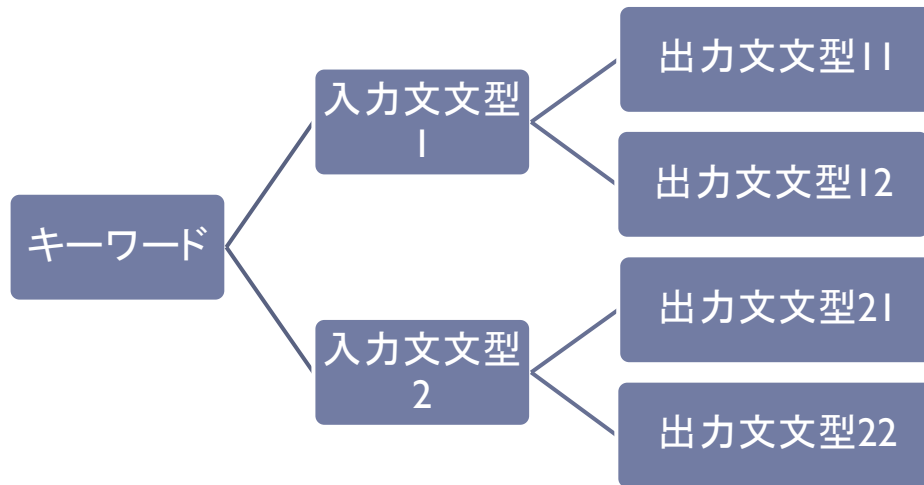
 (agentによって) object  verb される。



対話における文生成の例：ELIZA

- ・1966年発表
- ・精神科での患者と医師の対話をシミュレート
- ・デモ：emacsのdoctor

入力文を調べて、everybody, nobody, I, youなどのキーワードを探す。



マッチングする文型がないとき
Why do you ask?

キーワードがないとき
Please go on.
とか I see.

入力文文型：I am A.



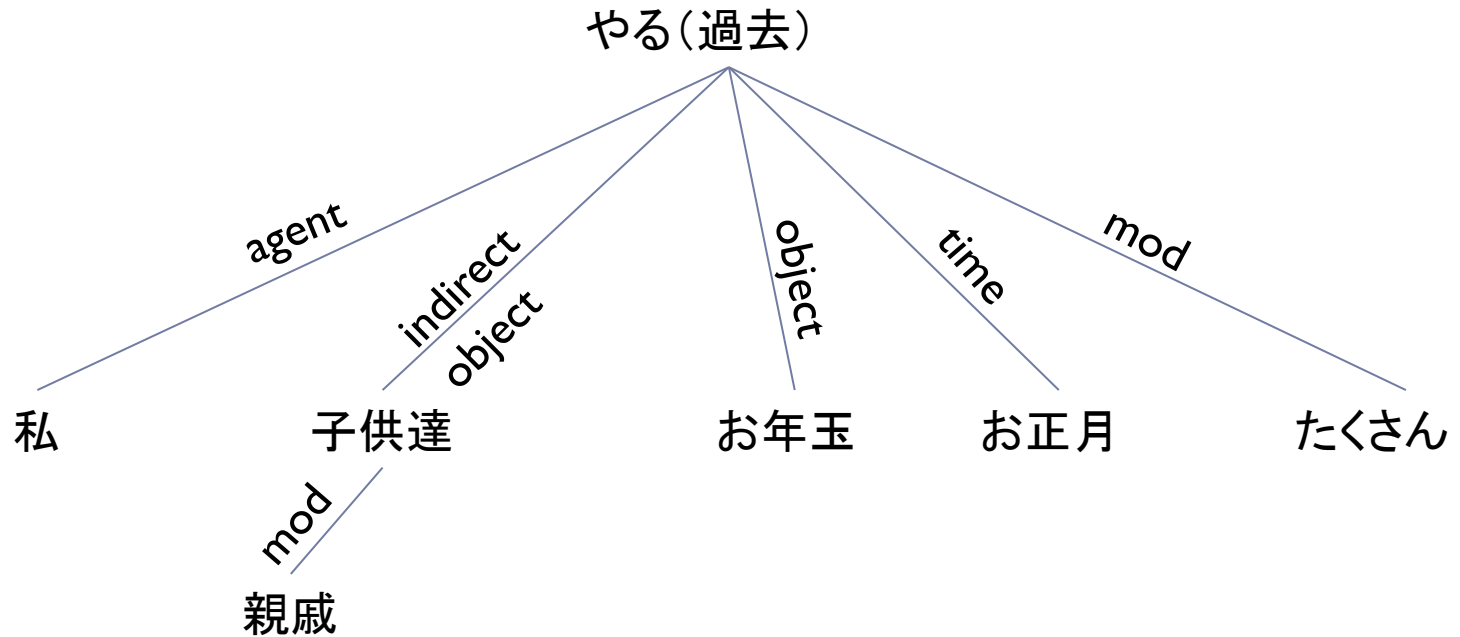
“I am helpful.”

出力文文型：what makes you think I am A.



“What makes you think I am very helpful.”

格構造からの文の生成



- ▶ お正月に私は親戚の子供達にお年玉をたくさんやった。
- ▶ 親戚の子供達はお正月に私からたくさんのお年玉をもらった。

文生成を決定する要因

表 9.3 発話の条件としての因子

	記号	
文 型	A	(1)平叙文, (2)比較文, (3)疑問文, (4)感嘆文, (5)条件文, (6)質問文
	B	(1)肯定文, (2)否定文
	C	(1)宣言, (2)叙述, (3)依頼, (4)願望, (5)義務, (6)可能, (7)当然, (8)仮定, (9)想像, (10)勧誘, (11)共同
	D	(1)現在, (2)過去, (3)未来, (4)無時制
	E	(1)default, (2)進行中, (3)繰り返し, (4)完了, (5)開始, (6)状態継続
	F	(1)絶対時間
発 話 者 の 態 度	I	(1)普通, (2)丁寧, (3)乱暴, (4)default
	J	話者が, (1)男, (2)女, (3)子供, (4)default
	K	聴者が, (1)男, (2)女, (3)子供, (4)default
	L	聴者に, (1)尊敬, (2)命令, (3)見下し, (4)対等, (5)default

因子に対応する動詞辞書

食事	標準	ご飯	「食べる」と連動する
	形式的	食事	「する」と連動する
	乱暴(JI,KI)	めし	「かう」と連動する
たべる	標準	たべる	
	乱暴(JI)	かう	
	I2	いただく	



文生成を決定する要因

「今日は天気がよい」 (A1,B1,C2,D1,E1)(I4,J4,K4,L5)

A1→A3, D1→D3

☞ 今日は天気がよいだろうか。

I4→I2

☞ 今日は天気がよいでしょうか？

最初の文 + C2→C4

☞ 今日は天気がよいとよいが！

さらにD1→D3

☞ 今日は天気が良くなるとよいが！

I4→I2

☞ 今日は天気が良くなるとよいのですが。

J4→J2

☞ 今日はお天気が良くなるとよいですね。



文生成における諸問題

- ▶ 標準的な語順
- ▶ 旧情報と新情報
- ▶ 指示詞
- ▶ 発話内容の表現と文章



標準的な語順

1. 広い範囲にかかる語は、そうでない語より先にくる。係り受けが非交差条件を満たすような構造となる。
2. 時格・場所格・主格・与格・対格・文末の順が標準的である。
3. 長い文節は短い文節の前にくる。
4. 文脈指示語を含む文節は、そうでない文節より前にくる。
5. 話題提示の句(～は、～も、など)は、文の先頭にくる。「～は」は、「～も」よりも前にくる。
6. 場所については、「～から～に(へ)」の語順が標準的である。
7. 動作述語と状態述語とではその左にくる文節の語順が次のようになる。
 - ▶ (+animateの語)が・(場所)に(で)・動作述語
 - ▶ (場所)に(で)・(-animateの語)が・状態述語
8. 「内容に関する部分文・相手・述語」という語順も多い。

旧情報と新情報

主題

彼はニューヨークに昨日行った。

題述

前提

焦点

主題

彼は昨日ニューヨークに行った。

題述

前提

焦点

文中の語順は、古い情報を表す要素から新しい情報を表す要素へと進み、最新の情報(焦点、focus)を表す名詞句は述語の直前に位置する。

英語では、文末に新情報が置かれるのが普通。

I found the book **at his desk**.

▶ The book was found at my desk **by himself**.

指示詞

太郎は₍₁₎太郎の₍₂₎本を太郎が₍₃₎先日会った太郎の₍₄₎女友達に与えた。

(3)主文の主語と一致なので省略

太郎は₍₁₎太郎の₍₂₎本を太郎が₍₃₎先日会った太郎の₍₄₎女友達に与えた。

(2)二個目の太郎を代名詞

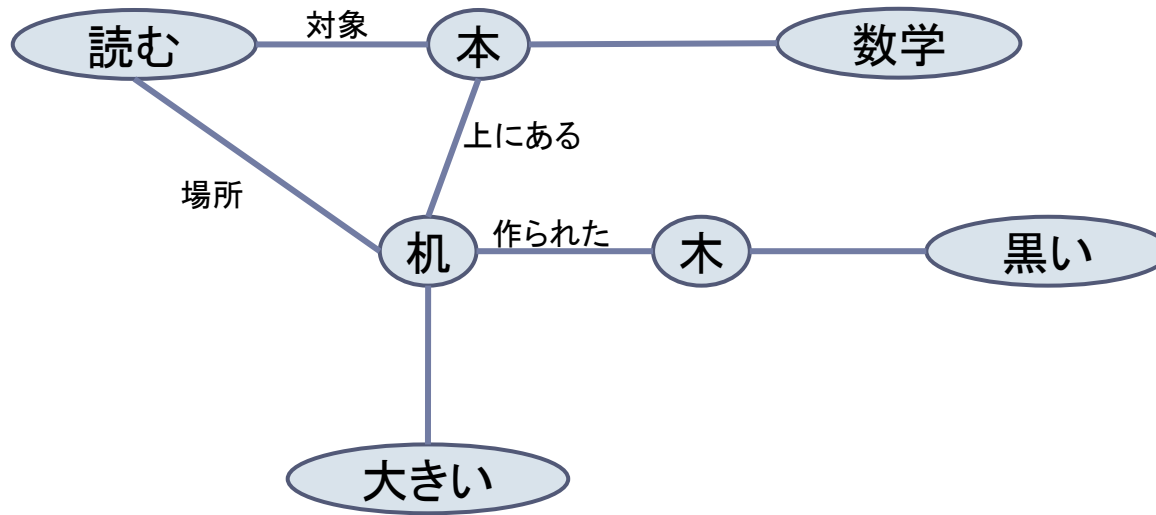
太郎は₍₁₎彼の₍₂₎本を太郎が₍₃₎先日会った太郎の₍₄₎女友達に与えた。

(4)変化

太郎は₍₁₎彼の₍₂₎本を太郎が₍₃₎先日会った自分の₍₄₎女友達に与えた。



発話内容の表現と文章



黒い木で作られた大きな机の上にある数学の本を読む。

OR

机で本を読む。その本は数学の本である。大きな机の上にある。それは黒い木で作られている。



文生成のステップ

1. 表現すべき内容を入力(格フレーム)
2. 表9.3の因子から対応する動詞を選択。
語順は、
場面、行為主体、目的、時間、場所、範囲、被益者、対象、行為が標準的
3. 発話者、聴者、文中の第三者の関係により、文中に現れる人物、文末表現の選択。
一貫した流れ(coherency. 結束構造)。
 1. 主題や焦点を決定。どちらかを主語にする。
 2. 文脈上明らかな対象は指示詞、または省略。
 3. 2回目の出現は、「自分」、「それ」、省略。
 4. 埋め込み文が長い場合は、二つの文に。



まとめ

- ▶ 簡単な文の生成法
 - ▶ 文のランダム生成
 - ▶ スロット法
 - ▶ 言い換え
 - ▶ 文の変形
 - ▶ 対話における文生成の例: ELIZA
 - ▶ 格構造からの文の生成
 - ▶ 文生成を決定する要因
 - ▶ 文生成における諸問題
 - ▶ 標準的な語順
 - ▶ 旧情報と新情報
 - ▶ 指示詞
 - ▶ 発話内容の表現と文章
 - ▶ 文生成のステップ
-



課題

▶ 文

「自然言語処理の授業で発表する。」

に対して、

1. 「文生成を決定する要因」に応じて、元の文を変形せよ。
 1. A6, B3, C3
 2. C6, D2, I2
2. 自分で文を作り、表9. 3により変形した文を3文示せ。その際、変形の因子を示すこと。