

平成 21 年度自然言語処理定期試験問題

実施日:平成 22 年 1 月 18 日(月)1限

試験会場:講実 A403 試験時間:60 分 担当:亀田弘之

問題0 学籍番号・氏名を所定の箇所に正しくかつ丁寧に書くこと。[10 点]**問題1** 言語 $L = \{ \text{Students read a book. Teachers ate a cake.} \}$ に関して、次の (1) ~ (2) に答えよ。[20 点]

- (1) 言語 L を定める文法 G を作成せよ。[10 点]
 なお、必要に応じて以下の用語を適宜使用しなさい。
 s : sentence (文) $subj$: subject (主語) obj : object (目的語)
 np : noun phrase (名詞句) vp : verb phrase (動詞句) v : verb (動詞)
 det : determiner (限定詞) art : 冠詞(article) n : noun (名詞)
 $prpn$: proper noun (固有名詞)
- (2) 上記 (1) で作成した文法 G に基づき、言語 L の文を対象とする構文解析プログラムを作成せよ。[10 点]

問題2 以下のような4種類の文書(文書1~文書4)がある。

以下の (1) ~ (2) に答えよ。[20 点]。

- (1) インデックスファイル表を作成せよ[5 点]。
 (2) TF-IDF 法の表を作成せよ。[10 点]
 (3) キーワード「言語」と「翻訳」の AND 検索を TF-IDF 法に基づいて行うとき、検索結果はどのようになるか。 [5 点]

事実:

表. 文書とそこに含まれるキーワード

文書名	含まれるキーワードとその出現回数(延べ)
文書1	言語(2)、コンピュータ(1)、翻訳(2)
文書2	コンピュータ(1)、翻訳(2)
文書3	言語(2)、翻訳(3)、知識(2)
文書4	翻訳(2)、知識(1)

(注1) キーワードの後のカッコ付き数字は、そのキーワードの延べでの出現回数を表している。

(注2) このような表は転置インデックス(索引)とも呼ばれる。

問題3 NLP(自然言語処理)技術を利用する新しいアプリケーションを1つ提案しなさい。その際、アプリケーション名、利益を受ける人、サービスを提供する人を明記するとともに、そのサービスがなぜ価値を持っているのか等を明瞭に述べよ。[20 点]**問題4** 次の用語のうち1つを選び、知るところを200文字程度で簡潔に述べよ[20 点]。

- (1)自然言語 (2)形態素 (3)形態素解析 (4)構文解析 (5)形式文法

問題5 日本語文「メロスは激怒した」に対して、以下のような構文解析プログラムを作成した。これを形態素解析もあわせて実行するように改良せよ[10 点]。

$\left\{ \begin{array}{l} \text{文}(A, C, \text{文}(_ \text{主語}, _ \text{述語})) :- \text{主語}(A, B, _ \text{主語}), \text{述語}(B, C, _ \text{述語}). \\ \text{主語}(A, C, \text{主語}(_ \text{名詞}, _ \text{助詞})) :- \text{名詞}(A, B, _ \text{名詞}), \text{助詞}(B, C, _ \text{助詞}). \\ \text{述語}(A, B, \text{述語}(_ \text{動詞})) :- \text{動詞}(A, B, _ \text{動詞}). \\ \text{名詞}([\text{メロス}], [], \text{名詞}(\text{メロス})). \\ \text{助詞}([\text{は}], [], \text{助詞}(\text{は})). \\ \text{動詞}([\text{激怒した}], [], \text{動詞}(\text{激怒した})). \end{array} \right.$

以上