【１】形式言語　　　　　　　　　　　　　　　　　　　　　　　　　　　　　…テキストP.（　　　）

日本語や英語のように人が生活していくなかで自然に生まれた言語を**自然言語**といい、特定の目的のために人工的に作成された言語を**形式言語**と呼ぶ。

１．正規表現

**正規表現**とは、いくつかのパターンの文字列を、一定の形式で表現する方法の１つである。通常の文字と、メタキャラと呼ばれる特別な意味を持つ記号を組み合わせて表記する。

代表的なメタキャラには次のようなものがある。

｜ 　　いずれかの文字を選択する場合に使われる。

“tax|tex”はtaxまたはtexのことである。

( ) 　　まとめて表現する場合に使われる。

まず( )の中を処理し、その後( )の外の文字列を連結する。

“t(a|e)x”では、“(a|e)”の部分でaまたはeを選び、その後tとxを連結する。

？＊＋　　同じ文字を繰り返す場合に使われる。

？は直前の文字が、０個または１個、＊は０個以上、＋は１個以上あることを表す。

≪範例１≫

UNIXにおける正規表現［A－Z］＋［0－9］＊が表現する文字列の集合の要素となるものはどれか。ここで，正規表現は次の規則に従う。

［A－Z］は，大文字の英字１文字を表す。

［0－9］は，数字１文字を表す。

＋は，直前の正規表現の１回以上の繰返しであることを表す。

＊は，直前の正規表現の０回以上の繰返しであることを表す。

ア　456789 イ　ABC＋99 ウ　ABC99＊ エ　ABCDEF

≪解答≫　エ

ア　４５６７８９

文字列の先頭に、少なくとも英字が１文字なくてはなりません。

イ　ＡＢＣ＋９９

この文字は表現できる文字に含まれません。

ウ　ＡＢＣ９９＊

この文字は表現できる文字に含まれません。