

NLP adalah mencoba untuk membuat komputer dapat mengerti perintah-perintah yang ditulis dalam standar bahasa manusia. NLP tidak memperdulikan bagaimana sebuah kalimat dimasukkan ke komputer tetapi menyalin informasi dari kalimat tersebut.

Pendekatan-pendekatan pada NLP: Inti dari NLP adalah PARSER, Dimana PARSER tersebut membaca setiap kalimat, kata demi kata, untuk menentukan apa yang dimaksud.

PARSER terdiri dari 3 jenis: PARSER STATE-MACHINE, PARSER CONTEXT-FREE RECURSIVE-DESCENT, PARSER NOISE-DISPOSAL.

Hal-hal yang bertentangan dengan pendekatan NLP -: Bahwa sesungguhnya NLP menggunakan semua informasi dalam sebuah kalimat, hanya manusia yang dapat melakukan hal tersebut. Mencoba memperbolehkan komputer menerima perintah bahasa alami, tetapi hanya menyalin inti informasi pada perintah (command).

Batasan Bahasa, Aspek yang paling sulit dalam pembentukan sistem pengendali NLP adalah: Pengakomodasian kekompleksan dan kefleksibelan bahasa manusia dalam sistem. Contoh Bentuk standar: Subyek-Verb-Obyek.

Diasumsikan: Adjective mengawali Noun, Adverb mengawali Verb, Semua kalimat diakhiri dengan titik.

Parser state-machine menggunakan keadaan yang sesungguhnya dari kalimat untuk memprediksi tipe apa dari kata yang berlaku.

State-machine: directed graph yang menunjukkan transisi yang valid dari satu state ke yang lain.

Kegunaan state-machine: a) dapat memotong atau memilah kalimat ke dalam komponen-komponennya; b) menentukan apakah sebuah kalimat dibentuk dengan benar dalam batasan dari grammar G1.

Database yang harus dibentuk sebelum implementasi: i) Membentuk kosa kata (vocabulary) dari kata-kata yang dikenal ke sistem dengan mengikuti tipe yang ada; ii) Menyimpan keadaan sesungguhnya dari kalimat.

Masalah yang paling buruk dengan state-machine parser adalah: Kekompleksannya; Parser tidak mengetahui bagaimana mencapai suatu keadaan.

Keuntungan parser state-machine : Ideal untuk aplikasi tertentu seperti: beberapa aplikasi database.

Contoh: sebuah kalimat adalah gabungan dari berbagai item dan item ini adalah gabungan dari item lain dan seterusnya sampai dipotong (dipilah) ke elemen-elemennya seperti Noun, Adjective, dan sebagainya.

Aturan-aturan yang ada pada setiap bagian yang telah dibentuk disebut: Production rule dari grammar.

Parser context-free menggunakan production rule untuk menganalisa sebuah kalimat.

Parser membentuk tipe parse tree yang disebut: 'context free', sebab Tree bukan dasar dari konteks setiap elemen. Hal ini berarti bahwa aturan atau rule akan bekerja untuk suatu statement yang menyerupai grammar G1 tanpa mengharapkan pada konteks setiap phrase.

Kegunaan context free selain untuk program-program AI NLP, juga untuk bahasa-bahasa komputer lainnya seperti: PASCAL, BASIC, C, MODULA-2.

Parser recursive-descent menggunakan: kumpulan rutin rekursif di mana menurunkannya melalui production rule sampai kalimat selesai ditelusuri seluruhnya.

Untuk membentuk parser context-free recursive-descent dibutuhkan beberapa vocabulary database dan dukungan predicate untuk menyalin kata-kata dari sebuah kalimat sebagai parser state-machine yang digunakan.

Keuntungan Parser Context-free recursive-descent: Mudah diimplementasikan dalam turbo prolog; Dapat berkomunikasi dengan kalimat baik tingkatan kata dan phrase; Mengetahui di mana parser dalam kalimat pada setiap saat.

Kerugiannya: Tidak dapat menangani cara valid dalam jumlah besar di mana kalimat dalam suatu bahasa dibentuk.

Tipe parser ini sesungguhnya sangat umum dalam aplikasi tipe database, seperti: Command processor.

Bahasa program di input, lalu dicek atau analyst wordnya ada/tidak di dictionary, sebelum dicek syntaknya, kemudian dicek syntaxnya dengan melihat grammarnya,

hasilnya berupa parse tree atau structure grammar, setelah itu dianalisis arti atau makna katanya, kemudian direpresentasikan dalam kode yang sudah tahu maunya.

Proses Penterjemahan dalam mengerti bahasa alami (Natural Language Understanding), ada 4 tahap: Lexical Analysis: cek masing-masing word & dictionary look up, Syntax Analysis (Parsing): sesuai grammar, Semantic Analysis: mengecek masing-masing arti kata, Discourse: Melihat semua kalimat yang lain, Sifatnya kompleks, Mengecek arti kata secara keseluruhan, Hukum/sesuai daerahnya.