Knowledge Representation for Natural Language Processing

Dan Tappan

Prepared in partial satisfaction of the requirements
for the doctoral comprehensive exam

Department of Computer Science
New Mexico State University

October 2001

Preface

Knowledge, in its many and varied forms, plays an indispensable role in the lives of every creature that interacts with the world. Its breadth and depth range from the massive amount of information humans rely on to the autonomous reactions of simple organisms. As we demand that computers play an ever-larger role in our world, they must somehow be embodied with an understanding of how this world is organized and how it functions. Mapping the world (or typically, an appropriate subset) to a form amenable to computer processing is the primary goal of knowledge representation [45:15,157; 2:391].

This paper focuses mainly on the knowledge required to process natural language in written form. However, the theory and practice extend equally well to many other applications with similar demands; e.g., artificial intelligence, speech recognition, robotics, computer vision, planning, etc. To this end, the framework is organized as follows:

Part I: Natural Language Knowledge

What kinds of natural-language knowledge are there?

Knowledge in natural language is used on multiple levels for a variety of purposes. Textbooks often generalize a hierarchy of linguistic building blocks, each of which contributes to the more complex levels above it. A similar structure is adopted here: at the lowest level is morphology, followed by syntax, grammar, semantics, pragmatics, discourse, and finally, at the highest level, world knowledge. While the field of knowledge representation generally focuses on the higher levels, it is argued here that all these levels have their own contributions of knowledge that need to be considered.

Part II: Knowledge Representation

What kinds of knowledge representations are there?

Knowledge can be used for many applications besides natural language processing. This part presents a survey of knowledge representation in theory and practice with the intent of solving issues from Part I. The devilish details of intersecting these parts will comprise a dissertation and the remainder of my career!

Contents

Part I

Natural Language Knowledge

Language is a complex puzzle of knowledge on various levels, as illustrated in Figure I.1. The organization and most appropriate presentation of these levels is subject to debate, and undoubtedly, no single approach can be considered truly correct. Part I of this paper presents a survey of language and linguistics based on the popular layered model [9:82; 35:599; 2:10]. Each section is built upon previous sections and addresses a set of relatively independent questions:

Section 1: Morphology
How are words formed?

Section 2: Syntax
What are the roles that words play?

Section 3: Grammar
How are words connected to form sentences?

Section 4: Semantics
What is the meaning of words and sentences without respect to any particular context?

Section 5: Pragmatics
What is the meaning of words and sentences in a specific context?

Section 6: Discourse
How are sentences organized to convey large-scale meaning?

Section 7: World knowledge
How does an understanding of the world contribute to language?

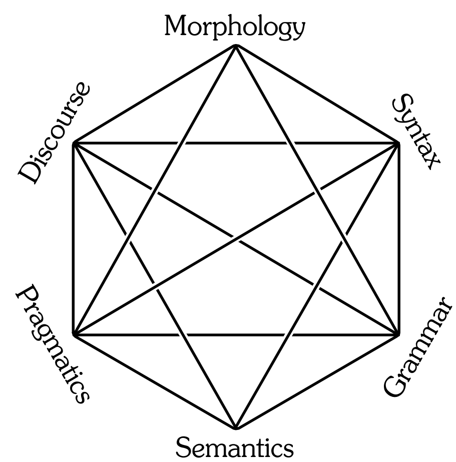


Figure I.1: *Linguistic Interrelations*

Knowledge of Language

The field of *knowledge representation* tends to focus on meaning—semantics, pragmatics, discourse, and world knowledge—more than it does on structure—morphology, syntax, and grammar [2:391]. This paper takes the view that all these levels play an essential role in language and must therefore be considered in the representation schemes to be discussed in Part II. Indeed, this holistic view of representation is shared by other authors [39:73]:

> Producers and consumers of text communicate successfully by manipulating knowledge that includes representations of various meanings of language elements, knowledge about the speech situation, including the knowledge about the interlocutor(s), knowledge about analysis and generation of the various language elements, and knowledge about the world in general.

Understanding language depends on various types of knowledge that overlap well with the structure of this paper; i.e., general knowledge about the fundamental components of language, the structure of coherent discourse, and the world, as well as specific knowledge about the current situation, the beliefs of the producers, and the beliefs of the consumers [45:716].

What is Language?

Language, the common thread throughout this paper, is perhaps best introduced by the properties that describe its various forms.

All communication systems, in particular the natural languages used by humans and the artificial languages used in knowledge representation, share three properties [27:20; 41:334]:

- *Mode of communication*

  Communication is transmitted in some way; e.g., vocally, visually, chemically, etc.

- *Semanticity*

  Communication uses meaningful signals.

- *Pragmatic function*

  Communication serves some useful purpose.

Some communication systems, including all natural and many artificial languages, exhibit some combination of more specific properties as well [27:20; 35:571; 57:131]:

- *Interchangeability*

  Messages can be both sent and received. Chemical systems of communication, as a counter-example, may leave only a meaningful trace to be interpreted but not reciprocated.

- *Cultural transmission*

  Language must be learned, even if genetic coding provides innate abilities.

- *Arbitrariness*

  Signals used for communication need not bear any resemblance to whatever they represent.

- *Discreteness or interdependence*

  Larger components of language are built from smaller components, at least down to some minimal units.

- *Displacement*

  Language can abstractly represent things that are not present or may not even exist.

- *Productivity, open-endedness, and creativity*

  Language is infinite in is ability to express anything.

The most powerful and widely used communication system, the natural languages of humans, builds even further upon these properties above [35:571; 57:131]:

- *Transmission through vocal-auditory channel*

  Sounds produced by the vocal organs are received by the ear.

- *Convertibility to other media*

  Language appears in written form, in visual form (e.g., sign language), and in tactile form (e.g., Braille).

- *Duality or double articulation*

  Language is composed of two layers: sounds or phonemes, which generally have no inherent meaning, contribute to the higher layer of morphemes, which do.

- *Continual change*

  Language continually evolves without losing its essential properties.

- *Turn-taking*

  Spoken language generally alternates between participants.

Section 1

Morphological Knowledge

***morphology**: …the internal structure and forms of words* [38]

For most people, the smallest unit of written language is the word. Although most words are actually built from lower-level components, the average language user typically is not aware of or does not consider anything beyond the visible surface form. The lower levels, however, must be addressed at some point in language processing. This subconscious understanding about the analysis and generation of word forms relies on morphological knowledge.

What is a Word?

Humans are undoubtedly deeply familiar with words in spoken form, and generally, in written form as well. Consequently, one would expect a straightforward definition of what a word is. Indeed, a native speaker of English may simply claim *anything with spaces around it*.[1] Such a simplistic definition works well for humans. It also illustrates why computers are confounded by natural language: humans excel at using it without really knowing or caring how it works.

To illustrate how such a definition is inadequate, consider the following five tests for determining word boundaries [9:91]. Each deals with separate nuances and idiosyncrasies of words, and none covers all cases. Furthermore, these tests depend on human intuition and judgment calls, something computers rarely emulate well.

- *Potential pause*

  When a sentence is spoken slowly, pauses tend to indicate word boundaries. However, speakers may also unintentionally emphasize *syllable* boundaries, as in *Chito / is / going / to / the / <u>mar / ket</u>*.

[1] Chinese speakers, on the other hand, would have to account for the lack of spaces in their language. Interestingly, although roughly 95 percent of Chinese words are composed of just one or two characters [18], agreement between native speakers on what constitutes a word is below 70 percent [52; 11:24].

- *Indivisibility*

  When new words are added to a sentence, they usually appear at the word boundaries, as in *Chito / is / <u>not</u> / going / <u>directly</u> / to / the / market*. However, this test fails to account for creative constructions like *unstinkingbelievable* or *craptacular*.

- *Minimal free forms*

  A word can be viewed as anything that can *meaningfully* stand on its own; e.g., *dog* or *large*, but not *the* or *after*.

- *Phonetic boundaries*

  In some languages, word boundaries are often indicated by stress or tone.

- *Semantic units*

  A word can be viewed as a semantic unit that includes collocated elements. For example, in *the dog <u>chewed up</u> the bone*, *up* would be considered part of *chewed*.

Even if these five tests can identify most words in a language, they provide little or no information as to how the words are related or used. Words are not simply defined by where they begin and end, but also by the roles they play. Even marginally different roles may be viewed quite differently in later stages of language processing. Nine categories of words are more or less common to all languages [35:1120]:

- *Orthographic word*

  Words as defined above in terms of spaces.

- *Phonological word*

  Words as they make sense in the flow of spoken language. For example, *a notion* and *an ocean* sound practically identical, and only context determines the intended meaning.

- *Morphological word*

  Words as they are legally constructed according to the spelling rules of a language.

- *Lexical word*

  Words grouped with their related forms. For example, *am*, *is*, and *are* are three different surface forms of the same verb *to be*.

- *Grammatical word*

  Words that define linguistic structure as opposed to content. For example, prepositions convey no meaning until they appear with nouns.

- *Onomastic word*

  Words typically considered *proper nouns*, such a *NMSU*, *Burger King*, *El Paso Diablos*, etc. Although such words are used frequently, dictionaries rarely include them.

- *Lexicographical word*

  Words with their related forms listed separately. This category is the opposite of lexical words.

- *Statistical word*

  Words considered in analytic terms, such as the number of nouns and verbs within a corpus, as well as how many of each particular noun and verb occur.

- *Other word*

  Words not falling into any of these categories. Interestingly, the bulk of these words describe other categories of words; e.g., buzzword, abbreviation, acronym, antonym.

How Languages Build Words

Most words are built from lower-level components. Their building blocks, known as *morphemes*, serve as the minimal units of meaning in any language [27:134]. While even lower-level components—*phonemes*, based on sounds—can be found, they do not systematically contribute to word structure [27:88].[2]

Some words, typically small, monosyllabic ones like *dog* and *cat*, are themselves morphemes because they have no subcomponents. Such words, known as *free morphemes*, are immediately usable with no manipulation required. The remaining words, known as *bound morphemes*, are exclusively subcomponents and do not become usable until they are appropriately combined with other morphemes; e.g., *pre–*, *–ed*, *–s*.

Combining morphemes

The morphological rules of a language specify how morphemes can be legally joined. For example, English freely permits the morpheme *anti–* to be added as a prefix to any noun, but it prohibits it as a suffix. Languages can be categorized as either *analytic* or *synthetic* according to their rules [27:64].

Analytic languages (also known as *isolating* or *root* languages) construct their meanings from free morphemes only. The morphological rules permit no affixes (i.e., prefixes and suffixes) and support very little or no manipulation of bound morphemes. Chinese is one such language: sentences are built by combining symbols, not by augmenting them (e.g., with new strokes) to vary their meaning or role [9:202].

Synthetic languages construct their meanings from both free and bound morphemes. The morphological rules vary considerably [27:166]:

- *Agglutinating language*

  Morphemes are combined to form larger meaning units in a way similar to how words are combined in English. The meaning of a single word is often based on the contents of its component slots. For example, verbs in Swahili are formed by filling the first slot with a bound morpheme for the subject pronoun, the second slot with a bound morpheme for the tense, and

[2] For example, the typical English vowel can be pronounced on average 20.7 different ways with little predictability [14]. For the purposes of this paper at least, it can be concluded that sounds are arbitrary and play no role in processing written language.

the final slot with a free morpheme for the verb stem. An English gloss of *u–li–soma*[3] appears as *you–PAST–write*.

The bound morphemes filling slots ordinarily have a single, unambiguous meaning, which makes dissecting the structure of a word relatively straightforward. In fact, the number of verb forms can be computed as the number of slots times the number of bound morphemes available to fill each.

- *Inflectional or fusional language*

  Morphemes are combined roughly in the same way as in agglutinating languages. The main difference lies in the nature of the bound morphemes, which do not have unambiguous meanings or clear boundaries. For example, in Russian, as in Swahili, verbs indicate tense and person with bound-morpheme affixes. However, the slots are not fixed, and the bound morphemes generally cannot be separated. An English gloss of □*ita–eš'* appears as *read–PRESENTyou*.

- *Polysynthetic or incorporating language*

  Morphemes are combined in radical ways to formulate entire sentences from single words. This process can be considered an extreme case of agglutination. Inuit (Eskimo) languages fall into this category.

How English Builds Words

English is an isolating, agglutinating, and inflectional language [9:295]. However, even with such varied morphological rules, it is comparatively weak in morphology! For example, English has at most five verb forms: *sing*, *sings*, *sang*, *sung*, *singing*. Italian and Spanish have around 50, whereas Modern Greek has roughly 350. Turkish has a whopping *two million* [41:127]!

Intra-word Constructions

Although English supports isolating and agglutinating morphological rules, the majority of words is formed according to inflectional rules by adding prefixes and suffixes [27:143]. The mechanism behind adding affixes depends on the type of construction, which may be either *inflectional* or *derivational*.

Inflectional constructions

Inflectional morphemes are also called *function morphemes* because they indicate the grammatical roles a word may play [27:136]. They exhibit the following properties in English [27:136]:

- They do not change the meaning of words or their grammatical role; e.g., *fast*, *fast–er*, and *fast–est* are just different degrees of the same property, and all are adjectives.
- They are required according to the rules of syntax; e.g., past-tense regular verbs take *–ed*.
- They produce many new words; e.g., adding *–s* nearly doubles the number of nouns.

[3] Hyphens are used here to indicate the morpheme boundaries. They are not present in normal writing.

- They appear after any derivational morphemes (see below); e.g., *immun–iz–ation–s*, where the plural morpheme –*s* appears last.
- They are suffixes only.

The complete set of English inflectional morphemes shown in Table 1.1 has remained stable since early modern English [27:135].

Morpheme	Meaning
–*s*	third-person singular
–*ed*	past tense
–*ing*	progressive
–*en*	past participle
–*s*	plural
–*'s*	possessive
–*er*	comparative
–*est*	superlative

Table 1.1: *English inflectional morphemes*

Derivational or lexical constructions

Derivational morphemes are also called *content morphemes* because they indicate the semantic roles a word may play [27:136]. They exhibit the following properties in English [27:135]:

- They change the meaning of words or their grammatical role; e.g., –*ly* generally transforms an adjective into an adverb.
- They are not required according to the rules of syntax; e.g., *un–tie* is legal with or without the prefix.
- They do not produce many new words because valid combinations are restrictive; e.g., –*hood* can be a suffix of only a small class of words like *brother* and *neighbor*.
- They appear before inflectional suffixes; e.g., *un–dress–ing*, where *un*– is a bound derivational morpheme, *dress* is a free morpheme, and –*ing* is a bound inflectional morpheme.
- They may be prefixes or suffixes.

An abridged set of derivational morphemes for the 386+ prefixes and 322+ suffixes in English is shown in Table 1.2 [10:198; 45:703; 23:242; [35:670].

Morpheme Class	Examples
abstract-noun makers	*–age*, *–dom*, *–ery*, *–ful*, *–hood*, *–ing*, *–ism*, *–ocracy*, *–ship*
concrete-noun makers	*–eer*, *–er*, *–ess*, *–ette*, *–let*, *–ling*, *–ster*
adjective-noun makers	*–ese*, *–(i)an*, *–ist*, *–ite*
adverb makers	*–ly*, *–ward(s)*, *–wise*
verb makers	*–ate*, *–en*, *–ify*, *–ize*
verb-to-noun makers	*–age*, *–al*, *–ant*, *–ation*, *–ee*, *–er*, *–ing*, *–ment*, *–or*
adjective-to-noun makers	*–ity*, *–ness*
noun-to-adjective makers	*–ed*, *–esque*, *–ful*, *–ic*, *–(i)al*, *–ish*, *–less*, *–ly*, *–ous*, *–y*
verb-to-adjective makers	*–able*, *–ive*

Table 1.2: *Sample English derivational morphemes*

Ambiguity

The great flexibility of English morphological rules provides a powerful mechanism for word building, but it also introduces problems. Morphemes can easily be combined to form constructions with uncertain or ambiguous structure. For example, Figures 1.1 and 1.2 show two conflicting interpretations of the adjective *un–lock–able* [27:141]. The first structure represents *can be unlocked*, whereas the second represents *cannot be locked*. The former is a *left-associative* structure because it combines the left-most morphemes at the lowest level; the latter combines the right-most morphemes into a *right-associative* structure. The correct interpretation depends on context, which will be discussed later in the sections on semantics and pragmatics.

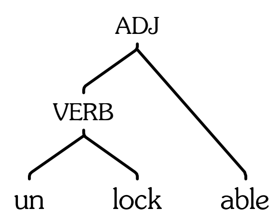


Figure 1.1: *Left-associative*

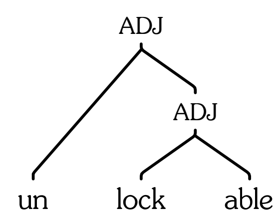


Figure 1.2: *Right-associative*

Extra-word Constructions

Just as morphemes can combine within a word to form a larger word, words can combine with other words to form larger units of tightly coupled meaning [9:91; 35:1120]. Indeed, the mechanisms and even the problems are quite similar. The remainder of this section considers several of the most common ways to combine words in English.

Multiword expressions

Although English is morphologically poor in terms of bound morphemes, it excels at forming new words by combining existing ones. This process, known as *compounding*, manifests itself in many ways that are beyond the scope of this paper. In general, however, they fall into three categories [35:244]:

- *Solid compounds*

  Words written together; e.g., *teacup*, *bookkeeper*, *steamship*. English generally does not build solid compounds with more than two or three words. Its sibling, German, on the other hand, frequently exceeds these limits, as in *Lebensversicherungsgesellschaftsangestellter* (life-insurance company employee) [45:704].[4]

- *Hyphenated compounds*

  Words linked with hyphens; e.g., *bridge-building*, *steel-lined*, *mega-lizard*.

- *Open compounds*

  Words written separately; e.g., *rocket launcher*, *operating system device driver*.

The rules for such combinations are inconsistent and often lead to more than one acceptable spelling; e.g., *businessman*, *business-man*, and *business man* [35:245]. Sometimes word stress can be used to indicate a distinction in meaning; e.g., *black bóard* versus *bláckboard* [27:144; 41:133].

The tree structures of compound nouns and morphemes are very similar. Not surprisingly, they share similar forms of morphological ambiguity, as illustrated by Figures 1.3 and 1.4.[5] The first structure represents *core of a fusion reactor*, whereas the second represents *reactor core made of steel*.

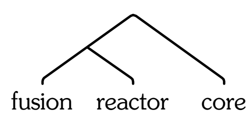


Figure 1.3: *Left-associative*

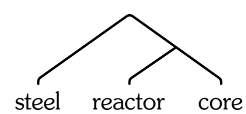


Figure 1.4: *Right-associative*

Even worse morphological ambiguity results from a combination of structures, as illustrated in Figure 1.5.[5] The intended reading is *contingency plan in case of a core meltdown in the fusion reactor*.

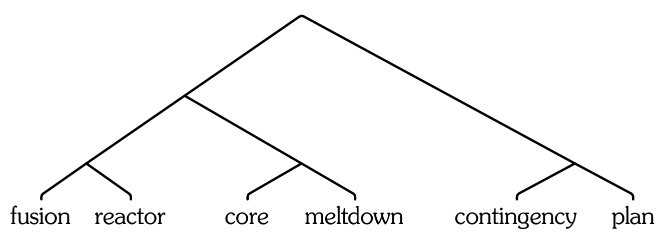


Figure 1.5: *Hybrid-associative*

[4] Although typically more restrained in noun compounding, German is fully capable of producing a behemoth like *Donaudampfschiffahrtsgesellschaftskapitänswitwenrentenauszahlungstag* (more or less: payday of retirement benefits to widows of captains of a steamship company providing trips on the Danube river) [59]!

[5] Figures 1.3, 1.4, and 1.5 are based on the examples in [39:22].

Finally, other common forms of multiword expressions may not lend themselves well to structural analysis [11:21]:

- *Foreign loan words*

  In expressions like *ad hoc*, *de facto*, and *quid pro quo*, perhaps the words could be considered individually in the source language, but in English, they should be taken as single units of meaning. The English pronoun *no one* also falls into this class.

- *Chained prepositions*

  Expressions like *out from under* and *in front of* should be considered single units of meaning, even though they contain multiple prepositions [62]. Some even have single-word equivalents that are favored; e.g., *despite* instead of *in spite of*.

- *Miscellaneous*

  Variations on dates, times, and numerical phrases are common. Identifying and processing them is usually dependent on the domain; e.g., military versus civilian dates and times.

Free-word combinations

Many multiword constructions are simply the result of logically combining words to form a new meaning; e.g., *the end of the road* [11:508]. The formula *end of* ____ can be filled by anything that reasonably has an end. Such combinations can be described by general rules and built according to the syntactic and semantic constraints on the headword [8]. It is not appropriate to treat them as single units of meaning because the overall meaning can be derived from their parts.

Idioms

For whatever reasons, seemingly meaningless constructions like *to kick the bucket* are sometimes accepted into general language use [11:508]. After the historical background is lost, a language is left with mysterious expressions that bear little or no semantic resemblance to whatever they were or are intended to mean. These expressions cannot be described by general rules and must simply be treated as single units of meaning. Morphological processing is generally limited, as they can rarely be manipulated beyond basic grammatical requirements (e.g., subject-verb agreement or tense).

Collocations

Words that habitually and systematically appear within the context of other words are usually considered collocates [2:310; 35:232]. They fall somewhere between free-word combinations and idioms, in that they are constructed according to general rules but are also restricted to certain word orders [11:508]. For example, the expression *table of contents* can also be a *table of figures*, *table of updated references*, etc. On the other hand, the expression *salt and pepper* appears awkward when written *pepper and salt*. While it is not strictly necessary to treat such expressions as single units of meaning, often some preferred ordering should be imposed.

Section 2

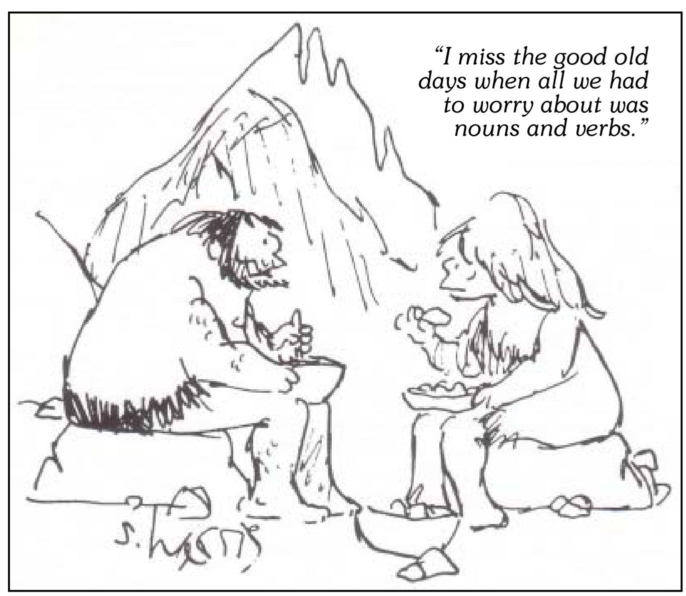
Syntactic Knowledge

syntax: *the arrangement of and relationships among words, phrases, and clauses forming sentences; sentence structure* [38]

The previous section discussed different views on what words are and how they can be constructed. However, it generally considered them only in terms of their form and content, not their function. This section extends the definition of words into the domain of what they do in language. Many textbooks treat syntax and grammar as a combined topic, but enough distinctions exist that separate sections are warranted in this paper.

What is Syntax?

Given any random collection of words, the average person should be able to separate them according to some commonsense "understanding" of their grammatical function; e.g., *these words are things, these words describe things, these words do stuff to things, these words connect things*, etc. Even such loosely defined groups are the foundation of syntax: each group does something different in language. Formal groups, known as *grammatical categories* or *parts of speech*, contain widely varying objects with possibly nothing at all in common but function. Legally combining different functions is the basis for building larger linguistic constructions, which will be discussed in the next section on grammar.



[9:99]

The Nuts and Bolts of Language

Pinker [41:106] writes: "A part of speech is not a kind of meaning; it is a kind of token that obeys certain formal rules, like a chess piece or a poker chip." Only the most important tokens—nouns, pronouns, adjectives, verbs, adverbs, prepositions, and phrase words—are considered here. The remaining tokens—determiners, conjunctions, and interjections—fit more appropriately into other discussions within this paper.

Nouns

The traditional grammar-school definition of a noun is a *person*, *place*, or *thing*. While it may adequately describe the physical world, it ignores much of the conceptual world; e.g., abstract qualities like *love*, actions like a *thud*, etc. In fact, English has six distinct classes of nouns [10:208; 53:19]:

- *Proper nouns*

  Instances of specific named people, places, objects, times, occasions, events, and so on. Such nouns are normally capitalized, can stand alone as clause elements (e.g., *Arizona is hot*), have no plural (e.g., **Arizonas*), and usually appear without determiners (e.g., **the Arizona*[6]).

- *Common nouns*

  Basically, any noun that is not proper. Such nouns are rarely capitalized, cannot stand alone as clause elements (e.g., **dog is hungry*), typically have a plural (e.g., *dogs*), and usually appear with a determiner (e.g., *the dog*).

- *Count nouns*

  Individual, countable objects. Such nouns allow a plural (e.g., *books*) and cannot stand alone in the singular (e.g., **book is long*) without a determiner (e.g., *a/the book*).

- *Noncount or mass nouns*

  Uncountable objects of mass or notion. Such nouns normally do not allow a plural (e.g., **butters*), can stand alone in the singular (e.g., *hockey is fun*), and may appear with certain indefinite determiners (e.g., *some advice*) or a definite determiner (e.g., *the music*).

- *Collective nouns*

  A countable variation of certain noncount nouns to group them as a collection of individuals or mass; e.g., *a blade of grass*, *a flock of sheep*, *a loaf of bread*.

- *Concrete nouns*

  Objects that can be observed and measured, such as *dog*, *book*, and *car*.

- *Abstract nouns*

  Objects that refer to an unobservable notion, such as *idea*, *thought*, and *love*.

[6] Example exceptions: *the Arizona I remember was not so hot* and *the Arizona*, referring to the battleship.

Figure 2.1 illustrates how these classes are related. Many nouns can belong to more than one class, and often the distinction is subtle [10:209].

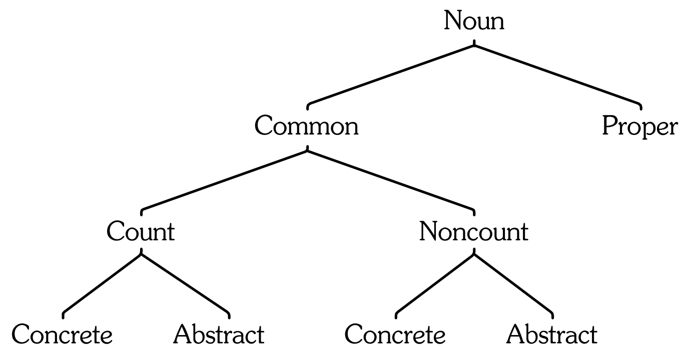


Figure 2.1: *English noun classes*

Noun Gender

Many languages assign an arbitrary grammatical gender to inanimate objects; e.g., a *knife* may be masculine, a *fork* feminine, and a *spoon* neuter. Animate objects have gender as well, but it tends to correspond with their natural gender; e.g., a *man* is masculine, and a *woman* is feminine. Gender in English is rudimentary and depends on the type of noun [10:209]:

- *Inanimate nouns*

  Nouns referring to objects that cannot be alive are usually 'it'; e.g., *car*, *table*, and *rock*.

- *Nonpersonal animate nouns*

  Nouns referring to living objects that are not people are usually 'it'; e.g., *tree*, *cockroach*, *wild animals*. A common exception is livestock, as shown in Table 4.2 on page 43.

- *Personal animate nouns*

  Nouns referring to living or dead people usually correspond to their real-world gender; e.g., *brother*, *sister*, *host*, *hostess*. Pet animals are often included in this class; e.g., *Spot*, *Tiger*, *Mr. Lizard*.

Pronouns

Nouns, as well as other larger elements, may be repeated within the same text. This redundancy adds to its length and decreases readability. Pronouns stand in for the repeated elements, as the following comparison illustrates:

> *My aunt Sarah saw the nice, old man wearing a green hat and asked the nice, old man wearing a green hat for the time. The nice, old man wearing a green hat told my aunt Sarah noon. My aunt Sarah thanked the nice, old man wearing a green hat.*
>
> *My aunt Sarah saw the nice, old man wearing a green hat and asked him for the time. He told her noon. She thanked him.*

Types of substitutions

English supports various forms of pronoun substitution [10:210]:

- *Replacing a noun without its modifiers*

  The pronoun refers to the noun only, not to any of its modifiers; e.g., *John has a new car, and Mary has an old one*.

- *Replacing a noun with some of its modifiers*

  The pronoun refers to the noun and some of its modifiers; e.g., *John has a new, fast car, and Mary has an old one*, means hers is old, but possibly also fast.

- *Replacing a noun with all of its modifiers*

  The pronoun refers to the entire noun phrase; e.g., *John has a new, fast car, and Mary has one, too*, means hers is also new and fast.

- *Referring to a general concept*

  The pronoun refers to one or more members of some class; e.g., *the dog saw someone on the road*, means an unspecified type of human, such as a man, woman, or child.

- *Referring to something unspecified*

  The pronoun refers to something known only in the context; e.g., *The man pointed to the sky and screamed, "Look at that!"* Here the pronoun refers to whatever he pointed to. Further context is necessary to clarify the reference; e.g., *"Its engine is on fire!"* would limit the object to being something that belongs in the air and has an engine, say an airplane. This form of pronoun usage depends heavily on *world knowledge*, which will be discussed later.

Types of pronouns

Since pronouns play varied roles, many types exist [10:210]:[7]

- *Personal pronoun*

  Refers to one or more specified objects; i.e., *I/me*, *you*, *he/him*, *she/her*, *it*, *we/us*, *they/them*.

- *Reflexive pronoun*

  Refers backward to one or more objects; i.e., *myself*, *yourself*, *himself*, *herself*, *itself*, *ourselves*, *themselves*.

- *Possessive pronoun*

  Refers to one or more objects in possession; i.e., *my/mine*, *your/yours*, *his*, *her/hers*, *its*, *our/ours*, *their/theirs*.

- *Reciprocal pronoun*

  Refers to a bi-directional relationship between objects; e.g., *each other*, *one another*.

[7] The discussion on anaphora (see page 63) considers pronoun usage in more detail.

- *Interrogative pronoun*

  Refers to objects in question; i.e., *who*, *whom*, *whose*, *which*, *what*.

- *Relative pronoun*

  Refers to objects being linked between different parts of a sentence; e.g., *the dog that barked was old* combines *the dog barked* and *the dog was old* into one sentence. This process will be covered next in the section on grammar.

- *Demonstrative pronoun*

  Refers to a contrast between near and far objects; e.g., <u>*this one*</u> *is broken; take* <u>*that one*</u>.

- *Indefinite pronoun*

  Refers to a notion of quantity; e.g., *somebody*, *anybody*, *everybody*, *nobody*. Also found in this class are items like *all*, *both*, *neither*, *none*, *each*, *much*, *many*, *more*, *most*, *less*, etc.

Adjectives

The physical and conceptual objects represented by nouns can have almost any properties (e.g., size, color, age, etc.). These distinctions are reflected in nouns by several types of adjectives that modify them [10:211]:

- *Attributive adjective*

  Adjectives that appear before the noun they modify; e.g., *the red car*.

- *Predicative adjective*

  Adjectives that appear after the noun they modify and certain *to be* verbs; e.g., *the car is red*.

- *Adverbial adjective*

  Adjectives that follow qualifying adverbs; e.g., *the very old car*.

- *Comparative adjective*

  Adjectives that compare the noun they modify with other related nouns; e.g., *the older car*.

How adjectives change the meanings of nouns will be discussed later in the section on semantics.

Verbs

Verbs describe what the physical and conceptual objects in the world can do and what can be done to them or with them, etc. [35:1083]. Parts of this discussion have been deferred until the next section on grammar, where they can be considered for their role in sentence structures.

Verb classes

Verbs indicate actions, describe states and conditions, and support other verbs [35:1083]. English has three types [10:212]:

- *Lexical or full verbs*

  This open class contains main verbs for every conceivable action; e.g., *run*, *jump*, *fly*, *eat*.

- *Modal verbs*

  This closed class contains auxiliary verbs that impart a degree of commitment onto the main verbs they modify. Only nine verbs are true modals: *can*, *could*, *may*, *might*, *will*, *would*, *shall*, *should*, and *must*; four others have similar enough function to be treated as modal as well: *dare*, *need*, *ought to*, and *used to*.

- *Primary verbs*

  Three verbs belong to both the lexical and modal classes: *be*, *have*, and *do*. Their function and interpretation depends on their usage; e.g., *John has a car* is lexical (in sense of possession), whereas *John has to buy a car* is modal (in the sense of necessity).

Verb forms

As illustrated in Table 1.1, English is morphologically very poor. In fact, verbs have only four distinct morphological forms, each of which is used in multiple ways to generate a wide range of English verb constructions [10:204; 2:28]:

- *Base or infinitive form*

  The dictionary form with no endings.

- *Simple-present form*

  The base form, for *I/you/we/they* subjects, or the base form with an *–(e)s* ending, for *he/she/it* subjects.

- *Present-participle form*

  The base form with an *–ing* ending, indicating a progressive action. Interestingly, the so-called *present* participle is not restricted to present-tense constructions; e.g., *the wind is/<u>was</u> blow<u>ing</u>*.

- *Past-participle form*

  The base form with an *–ed* ending (occasionally *–t*, for dialectal reasons, or *–en*, for irregular verbs), indicating one of the following:

  - to express a past aspect; e.g., *the man kicked the cat*
  - to express a passive construction; e.g., *the cat was kicked by the man*
  - to support certain clauses; e.g., *kicked across the room, the cat became angry*
  - to function as an adjective; e.g., *the injured cat*

Interestingly, the so-called *past* participle is not restricted to past-tense constructions; e.g., *the flag will be raised*.

A second overlapping distinction between verb forms is based on how verbs are grammatically associated with their subjects [10:212]:

- *Finite form*

  A verb limited by grammatical agreement to a particular person, number, tense, or mood; e.g., *he jumps, we jumped, they were*[8] *jumping*.

- *Nonfinite form*

  A verb not limited in the way a finite form is; e.g., *we have jumped, they were jumping, he might jump, she wants to jump*.

Verb tenses

English verbs can convey at least 16 distinct variations[9] on when an event occurs relative to other events, as well as whether the event completes. The details and nuances of these constructions are beyond the scope of this paper. In place of such a discussion, Figure 2.2 provides a graphical overview [2:28].

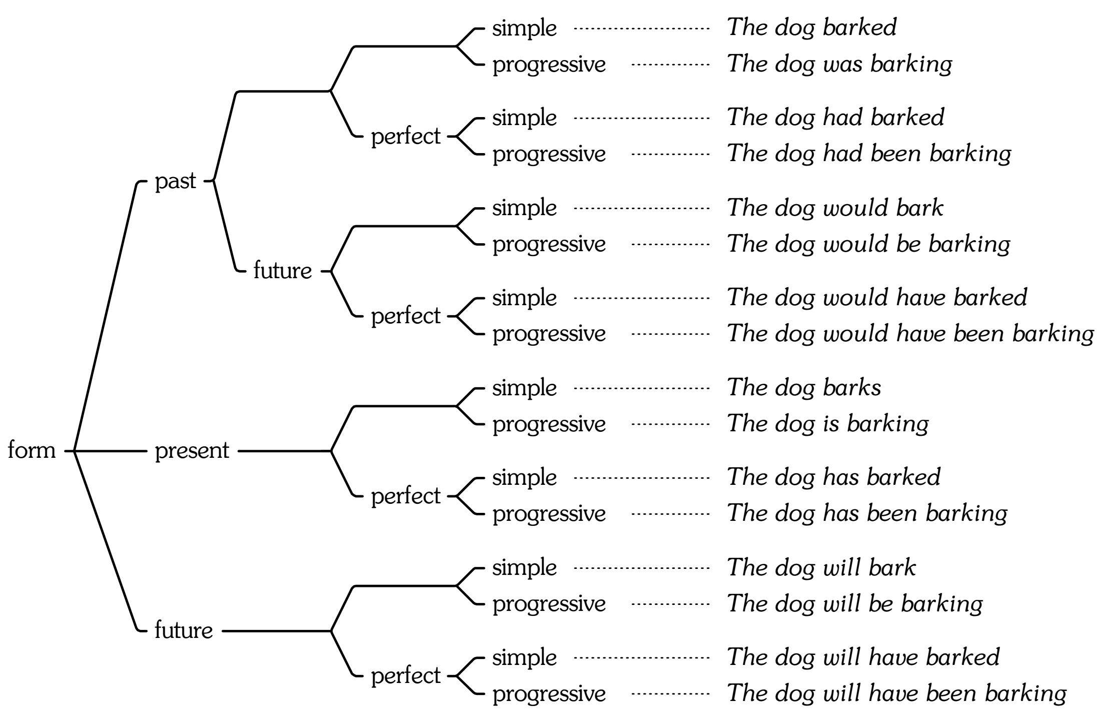


Figure 2.2: *English primary verb tenses*

[8] Note that the finite form is on the auxiliary verb when one is present, not on the main verb.

[9] A main verb may be preceded by up to four auxiliary verbs to generate more constructions than are considered here, but such usage is exceedingly rare; e.g., *the story must have been being written during the storm*.

Adverbs

Similar to the way adjectives modify nouns, adverbs modify adjectives and verbs. A distinction is made between adverbs that can modify either part of speech (e.g., *a somehow peculiar fellow* and *to succeed somehow*) and those that are specific to just one (e.g., *very quick* but not **to love very*; *to answer promptly* but not **a promptly correct answer*) [35:15].

Adverbs modify language in a wide range of subtle ways, as exemplified in Table 2.1 [2:35; 10:211].

Adverb Type	Examples
Degree	*very*, *rather*, *quite*, *too*
Location	*here*, *there*, *everywhere*, *nowhere*
Direction	*upward*, *downward*, *clockwise*
Manner	*slowly*, *quickly*, *fast*, *poorly*
Time	*yesterday*, *today*, *now*, *tomorrow*
Frequency	*always*, *sometimes*, *rarely*, *never*
Miscellaneous	*just*, *soon*, *somehow*, *whereby*

Table 2.1: *Sample adverb types*

Prepositions

If nouns, pronouns, adjectives, verbs, and adverbs are the building blocks of sentences, then prepositions are the glue that holds them together. The role and meaning of prepositions vary between simple and quite complex [35:801; 10:213]:

- *Simple prepositions*

  Single words usually specifying one relation; e.g., *at*, *about*, *before*, *after*, *above*, *below*.

- *Compound prepositions*

  Two prepositions usually specifying one relation; e.g., *because of*, *ahead of*, *due to*.

- *Complex prepositions*

  Multiple prepositions specifying a wide range of sequential relations; e.g., *in accordance with*, *on behalf of*, *as far as*, *out from underneath*, *up to but not including*.

- *Phrasal prepositions*

  Phrases covering aspects from Table 2.1; e.g., *before breakfast*, *every hour on the hour*, *for two hours*, *roughly six feet*, *in such a way*. This class overlaps considerably with adverbial phrases discussed in the next section on grammar.

A common usage of verbs, adverbs, and prepositions together gives English one of its most distinctive features and causes possibly some of the greatest difficulties for nonnative speakers [10:212; 35:772]. These structures, known as *phrasal verbs*, blur the line between literal, figurative, and idiomatic interpretations of their components. For example, verbs with clear movement like *to go*, *to put*, and *to take* often use adverbial particles like *up*, *down*, and *off* to indicate or emphasize motion: *to take* *something off a table*, *to put something up on a shelf or down on a chair*, and *to go off to the store*. In context with other verbs, however, they take on entirely different meanings; e.g., *to act up*, *to slow down*, and *to mouth off*. Even a superficial survey of this topic is well beyond the scope of this paper.

Phrase Words

Phrases in their wide range of forms comprise a significant portion of English text. Since they generally function as regular syntactic units do, they can be substituted into almost any valid context where conventional parts of speech would fit [35:776]:

- *Attributive phrases*

  Predicate phrases used attributively; e.g., *the tool was state of the art* becomes *the state-of-the-art tool*. So many variations of theses phrases are common in English that they are discussed separately below.

- *Nouns followed by prepositions and other nouns*

  Phrases modeled after certain French compounds; e.g., *tug-of-war*, *commander-in-chief*.

- *Phrases turned into lexical bases by addition of suffixes*

  Phrases treated roughly as nouns and modified according to normal morphological rules; e.g., *never-say-die-ism*, *state-of-the-artistry*.

- *Stunt formations*

  Phrases designed to be catchy and related to the topic; e.g., *whodunit* for a murder mystery.

- *Vague words*

  Phrases used in place of more appropriate words; e.g., *thingamajig*, *whatchamacallit*.

- *Phrasal verbs with noun derivatives*

  Phrases based on phrasal verbs, as discussed in the previous section.

Attributive phrases

As mentioned above, nearly any phrasal expression can modify nouns in English [35:776]:

- *Noun-based phrase*

  six-four-three double play, '*Freud and your mother*' *theory*, *group of friends get-together*

- *Verb-based phrase*

  Infinitive: *ready-to-wear clothes*, *time-to-kill rage*

  Participial: '*Contract with America*' *pledge*, *farmer-turned-astronaut story*

  Modal: *must-win game*, *can-do attitude, will-do-my-best attempt*

- *Adjective-based phrase*

  '*good news, bad news*' *joke*, *larger-than-life picture*, *holier-than-thou attitude*

- *Number-based phrase*
  thirty-pound weight, *2001 Jeep Cherokee*, *24/7 Internet access*
- *Preposition-based phrase*
  Preposition first: *behind-the-scenes information*, *over-the-counter medication*
  Preposition midway: *signal-to-noise ratio*, *rags-to-riches story*, *made-for-TV movie*
- *Coordination-based phrase*
  give-and-take politics, *hit-and-run driver*, *life-or-death decision*
- *Wh-words-based phrase*
  you-know-who, *what-the-hell-happened-here reaction*, *how-to book*
- *Negation-based phrase*
  no-man's land, *no-win situation, no-can-do attitude*

Section 3

Grammatical Knowledge

> ***grammar**: a body of rules imposed on a given language for speaking and writing it…* [38]

The function of key syntactic units was discussed in the previous section, but only in terms of isolated words or small constructions. For these units to form larger structures—clauses, phrases, and sentences—another, more advanced level of rules is needed. This section discusses a wide cross-section of aspects pertaining to grammar, which embodies such rules.

What is Grammar?

In many respects, defining what grammar is proves as difficult as defining the grammar of a language itself. The broad scope of what grammar allows, disallows, supports, denotes, emphasizes, and so on defies any simple description. The extent of this complexity is expressed well in the opening paragraph to a section on grammar in [9:88]:

> It is difficult to capture the central role played by grammar in the structure of language, other than by using a metaphor such as 'framework' or 'skeleton'. But no physical metaphor can express satisfactorily the multifarious kinds of formal patterning and abstract relationship that are brought to light in a grammatical analysis.

Pinker [41:334] describes grammar as a "discrete combinatorial system" with the following properties:

- *Infinite*

  There is no limit to the number of complex words or sentences in a language.

- *Digital*

  Infinity is achieved by rearranging discrete elements in particular orders and combinations, not by varying some signal along a continuum like the mercury in a thermometer.

- *Compositional*

  Each of the infinite combinations has a different meaning predicable from the meanings of its parts and the rules and principles arranging them.

Approaches to Grammar

In simplest terms, a grammar is a compilation of rules and guidelines for using language. The nature and presentation of the content vary according to the purpose and intended audience [9:88,97,413; [27:12; 41:371]:

- *Prescriptive grammar*

  Focuses on controversial constructions and specifies rules governing socially correct usage. Rules of this type focus on *linguistic performance*, or how language *should* be used.

- *Descriptive grammar*

  Describes grammatical constructions in a language without making evaluative judgments about their standing in society. Rules of this type focus on *linguistic competence*, or how language is *actually* used.

- *Pedagogical grammar*

  Supports teaching a foreign language or developing an awareness of a native language.

- *Reference grammar*

  Serves as a comprehensive reference book of grammatical facts and details. Quirk, et al. [43], with its 1,779 detailed pages, is the definitive source for English, and probably the largest resource ever written for any language.

- *Theoretical grammar*

  Investigates the commonalities and universal constructs of all human languages.

- *Traditional grammar*

  Summarizes attitudes, beliefs, and methods for grammatical study before formal linguistic methodology was established.

Different linguistic schools of thought impose various frameworks on analyzing and describing grammar. The following is a cross-section of popular approaches from the 20th century [9:412; 35:433]:

- *Functional sentence perspective*

  Analyzes utterances in terms of their information content.

- *Dependency grammar*

  Explains formal grammatical relationships by establishing dependencies (or valencies) between elements. See *Structural requirements* below for more details.

- *Tagmemics*

  Relates linguistic forms and functions.

- *Stratificational grammar*

  Defines language as a system of related layers, or strata, of structure. A similar approach is taken in this paper to describe natural-language knowledge.

- *Systemic linguistics*

  Treats grammar as a complex network of interrelated systems, especially for semantic and pragmatic analysis.

- *Generative grammar*

  Specifies precise, formal requirements that any structure must satisfy to be considered grammatically correct.

The generative-grammar approach, developed by Chomsky starting in the late 1950's, spawned numerous theoretical offspring [9:413; 2:131; 54:118; 45:686; 51:293]:

- *Case grammar*

  Examines the semantic roles (or cases) played by elements of sentence structure.

- *Relational grammar*

  Views grammatical relations (e.g., subject, object) as more important than the formal categories (e.g., noun phrases, verb phrase) of generative theory.

- *X-bar theory*

  Considers an alternative account of phrase structure within a generative grammar.

- *Montague grammar*

  Focuses on logical languages through a close correspondence between syntax and semantics.

- *Generalized phrase-structure grammar*

  Ignores the role of transformations in a generative grammar by providing an alternative account of phrase structures for grammatical analysis.

- *Functional grammar*

  Seeks alternatives to an abstract, formal approach to grammar, particularly a pragmatic view of language as social interaction.

- *Realistic grammar*

  Claims grammatical analyses should be grounded in the functionality of the human mind.

- *Network grammar*

  Tries to simulate how people understand language for applications in artificial intelligence.

- *Government-and-binding theory*

  Investigates the conditions and structural contexts that formally relate elements of a sentence.

Grammatical Structures

Language is compositional, which means its larger structures are built from smaller structures, and these structures, in turn, are built from even smaller structures, and so on, down to atomic morphemes [27:34; 9:95; 10:217; 56:419]. Such is the organization of this paper as well: each level of knowledge—and correspondingly, each section—contributes to subsequent levels. A graphical representation of this hierarchical structure is shown in Figure 3.1.

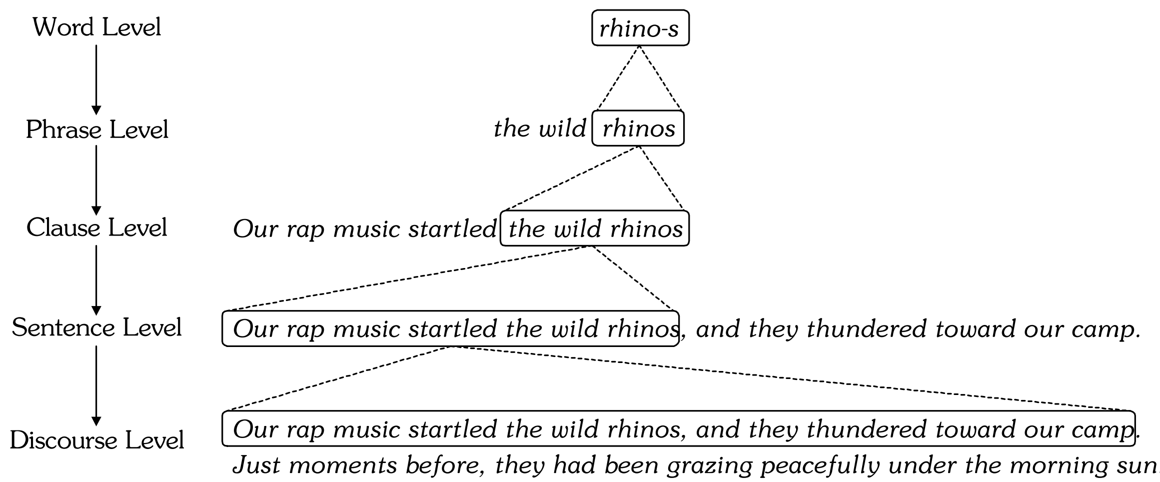


Figure 3.1: *Hierarchical representation of grammatical structures*

Structural requirements

Any hierarchy is structured according to certain organizational rules. Grammatical structures, in particular, adhere to three main principles that dictate how they can be built and joined [23:303; 53:106]:

- *Valency*

  A system of constraints adapted from chemistry in which interlocking elements must join or cancel each other for an expression to be syntactically and semantically complete. This system is similar in principle to subcategorization frames and selectional restrictions [11:567; 23:317]; e.g., the verb *give* has two primary frames [32:45,138; 45:670]:

  SUBJECT + *give* + DIRECT OBJECT + *to* + INDIRECT OBJECT
  SUBJECT + *give* + INDIRECT OBJECT + DIRECT OBJECT

- *Agreement*

  Grammar may be figuratively treated as a puzzle, where pieces can be joined only if they satisfy the constraints; e.g., subject-verb agreement in person and number.

- *Word order*

  English is highly dependent on word order to indicate grammatical interrelations because it has so few morphological markers; e.g., *the dog bit the man* is quite different from *the man bit the dog*.

Valency and agreement impose constraints on at least the following grammatical properties [9:93; 10:224; 23:304,307]:

- *Aspect*

  Conveys time-related details like duration, completeness, habitualness, continuousness, progressiveness, etc.; e.g., *John drove to the work* versus *John is driving to work*.

- *Case*

  Conveys syntactic cues like subject, direct and indirect object, possession, instrument, etc.; e.g., *John*$_{\text{SUBJECT}}$ *gave Mary*$_{\text{INDIRECT-OBJECT}}$ [*the book*]$_{\text{DIRECT-OBJECT}}$. English is considered by most to have only two cases: *subject* and *object* (e.g., *he* and *him*, respectively). Finnish, arguably the most extreme example, has 15 [9:92]!

- *Gender*

  Conveys either the natural gender for most living things or often an arbitrary gender for nonliving things. See *Noun Gender* on page 15 for more information.

- *Mood*

  Conveys factuality, possibility, uncertainty, likelihood, etc.; e.g., *If today were Saturday, we would not be working*.

- *Number*

  Conveys the quantity of a noun or nouns.

- *Person*

  Conveys the speaker, addressee, third party, etc.; e.g., *I am running for President*, as spoken by the candidate.[10]

- *Tense*

  Conveys the time when something happened or when it was reported.

- *Voice*

  Conveys what acted on what; e.g., *John stole the car* versus *the car was stolen by John* versus *the car was stolen*.

[10] The notable exception is Bob Dole, who consistently referred to himself in the third person during his 1996 campaign; e.g., *Bob Dole wants* [read: *I want*] *your vote*. A computer may conclude that the speaker and Bob Dole are not the same person.

Phrases

As Figure 3.1 illustrates, the hierarchical nature of language builds its larger structures from smaller ones. Phrases constitute the major building blocks in this process by systematically combining isolated words into semi-meaningful fragments. The rules behind constructing phrases are based on four properties [27:193]:

- *Generativity*

  An infinite set of grammatical constructions can be produced from a finite set of rules.

- *Infinite recursion*

  Production rules are composed of rules circularly composed of themselves and/or other rules.

- *Hierarchical structure*

  An independent, lower level of structure is built for each recursive call, and this structure contributes to the structure above it.

- *Ambiguity*

  Inherent ambiguity in language can be reflected in the grammar.

Phrase structures

Phrases structures are the result of applying phrase rules to individual words to build the major components of a sentence [10:222,227]:

- *Noun phrase*

  In grossly simplistic terms, a string of words headed by a noun. The variations on this theme are so great that noun phrases are discussed separately below.

- *Pronoun phrase*

  A small set of constructions headed by a pronoun; e.g., *nearly everyone*, *she herself*.

- *Adjective phrase*

  Usually an intensifier followed by one or more adjectives; e.g., *very big*, *quite big and fast*.

- *Verb phrase*

  Up to four auxiliary verbs followed by a main verb; e.g., *will have been barking*. See *Verbs* on page 17 for details.

- *Adverb phrase*

  Usually an intensifier followed by one or more adverbs; e.g., *quite often*, *very quickly indeed*.

- *Prepositional phrase*

  A preposition followed by a noun phrase; e.g., *in the old house*, *under the bed*.

More on noun phrases

Regardless of their complexity, all noun phrases can be described in terms of four components [10:222]:

- *Head*
  The object, typically a noun, to which the phrase refers; e.g., *the large* [*dog*].
- *Determiner*
  An indicator of appropriate semantic interpretation; e.g., *the*, *a*, *all*, *every*, *some*, *any*, *no*.
- *Premodifier*
  Anything between the determiner and the head; e.g., *the* [*extremely large and hungry*] *dogs*.
- *Postmodifier*
  Anything after the head; e.g., *the dogs* [*playing in the open field*].

Clauses

When phrases are joined properly, they form larger grammatical units called *clauses*, which, in turn, can be combined in various ways to build sentences [10:220; 9:95].

Clause types

How clauses are combined depends on their structure and the constraints imposed by the conjunctions that join them [35:220,638,263,998,1005]:

- *Main or principal clause*
  Any clause or sentence that can normally stand on its own; e.g., *John went to the restaurant*. See *Sentence structures* below.
- *Coordinate or independent clause*
  Two main clauses joined by a coordinating conjunction; e.g., *John ate fish* and *John drank beer* can be written as *John ate fish and drank beer*. See *Clause combinations* below.
- *Subordinate or dependent clause*
  Any clause that cannot normally stand on its own because of a subordinating conjunction; e.g., *…<u>since</u> it was raining*, *…<u>while</u> he was watching television*. To form a complete sentence, such clauses must be joined with a main clause. See *Clause combinations* below.
- *Superordinate clause*
  Any clause that contains another clause; e.g., *John ate fish because he was hungry*.

These four types of clauses can be further analyzed by how their verbs are formed [35:220,752,1086]. See *Verb Forms* on page 18 for more details.

- *Finite or participial clause*

  Any clause having a formed verb; e.g., *the dog barked*, or *until the war ended*, or *having completed his homework, Johnny went to bed* (where the subject of *having* is *Johnny*). Compare with *Absolute clause* below.

- *Nonfinite or infinitive clause*

  Any clause having an unformed verb; e.g., *to tell the truth*.

- *Verbless clause*

  Any clause having no explicit verb or subject; e.g., *if* [*it is*] *necessary, press* HELP.

Finally, two more types are based on how clauses are embedded in or linked with other clauses [35:220,6,14]:

- *Relative or adjective clause*

  Clauses that modify nouns through the direct or implied use of *who*, *which*, or *that*; e.g., *the man <u>who ate the apple</u> said it was tasty*, or *the apple* [*that*] *he ate was tasty*.

- *Absolute or adverb clause*

  Similar to a finite clause, but with independent subjects; cf. *the homework having been completed, Johnny went to bed* (where the subject of *having* is *the homework*).

Clause elements

These clause types are composed of subcomponents, or elements, that express a particular kind of meaning. Traditional grammar considers only the *subject* (basically whatever comes before the verb) and the *predicate* (everything else) [10:220]. A more detailed analysis recognizes five elements [10:220; 9:95]:

- *Subject*

  The theme or topic of the sentence, usually appearing in the first position.

- *Verb*

  The action being performed, typically by the subject.

- *Object*

  The element being affected by the action.

- *Complement*

  Essential information elaborating on the preceding clause.

- *Adverbial*

  Generally less important or optional information indicating time, place, manner, etc.

English is considered a subject-verb-object (SVO) language, which indicates the customary ordering of its clause elements. The remaining five orderings (i.e., SOV, OVS, VSO, VOS, OSV) are possible in other languages. Table 3.1 lists the basic element variations in English [10:220; 9:95].

Element Ordering	Example
Subject + Verb	*The dog + barks*
Subject + Verb + Object	*The dog + chased + the cat*
Subject + Verb + Complement	*The dog + is + hungry*
Subject + Verb + Adverbial	*The dog + went + under the table*
Subject + Verb + Object + Object	*The dog + gave + his trainer + the ball*
Subject + Verb + Object + Complement	*The dog + paid + his trainer + no attention*
Subject + Verb + Object + Adverbial	*The dog + chased + the cat + yesterday*

Table 3.1: *Common English clause orderings*

Clause combinations

Clauses can be joined in various ways by conjunctions [10:213; 9:95]:

- *Coordinating conjunction*

  Joins clauses of similar grammatical status, typically noun phrases or adjectives; e.g., *John was hungry <u>and</u> thirsty*. Members of this class include *and*, *or*, *but*, *either … or*, *neither … nor*, etc.

- *Subordinating conjunction*

  Joins clauses of different grammatical status, typically main and subordinate clauses; e.g., [*John went to the store*]$_{\text{MAIN}}$ [*<u>because</u> he needed milk*]$_{\text{SUBORDINATE}}$. Over a dozen subordinate meanings can be expressed; e.g., *since*, *until*, *because*, *if*, *whether*, *where*, *in order to*, *although*, etc.

Sentences

In simplistic terms, when phrases and clauses are combined legally, they form sentences. However, providing a solid formal definition of what really comprises a sentence turns out to be far more difficult [35:918]. This issue is better addressed within the domain of psycholinguistics and philosophy, as it relates closely with the notions of thought and intent. Only the grammatical aspects of sentences are considered here.

Sentence types

Different sentence types communicate varied information or the desire for information [35:918; 10:219]:

- *Declarative sentence*

  Statements generally conveying information; e.g., *the train is here*, or *yesterday it rained almost all day, but finally the sun came out*. The vast majority of non-narrative written sentences is declarative.

- *Interrogative sentence*

  Questions expecting a variety of responses:
  - *yes-no*, asking for an affirmative or negative reply; e.g., *are you tired?*
  - *wh*, asking for a reply in terms of *who*, *what*, *where*, *when*, *why*, or *how*
  - *alternative*, asking for a reply from a set of possibilities; e.g., *are you hot or cold?*
  - *exclamatory*, asking indirectly for a response to a claim; e.g., *is she fat or what?*
  - *rhetorical*, not really asking for a response at all; e.g., *who cares?*
  - *tag*, asking indirectly whether the listener is paying attention; e.g., *Canada is cold, eh?*
  - *echo*, asking for missed details in the previous statement; e.g., *Put it here. Put it where?*

- *Imperative sentence*

  Directives telling someone to do something; e.g., commanding, inviting, warning, pleading, suggesting, advising, instructing, permitting, requesting, meditating, expressing good wishes, expressing an imprecation (*go to hell*!).

- *Exclamatory sentence*

  Short sentences responding to something emotive; e.g., *Oh dear! What a mess! You jerk!*

Sentence structures

Many children start learning to read English with books that illustrate basic sentence structures [9:89]:

> I see Spot. Spot is a dog. Spot belongs to my neighbor. Spot is running. Spot runs fast. Spot is chasing a ball. The ball belongs to Johnny.

As their reading and writing skills improve, children increase the complexity of their constructions. Eventually, by adulthood[11], they have a large repertoire of powerful ways to combine and express their ideas grammatically; e.g., *I see Spot, my neighbor's dog, running fast after Johnny's ball.*

What differentiates simple sentence structures from advanced ones is how their clauses and phrases are combined [35:918,936,244,243]:

- *Simple sentence*

  One main clause, no coordinating or subordinate clauses; e.g., *I see spot.*

- *Compound sentence*

  Two or more main clauses linked by some coordinating conjunction like *and* or *however*, or by a semicolon; e.g., *Spot is chasing a ball, and the ball belongs to Johnny.*

- *Complex sentence*

  One main clause with one or more embedded subordinate clauses; e.g., *I see Spot, who is my neighbor's dog.*

[11] At least, one would hope. Unfortunately, *Spot*-like prose appears all too common among adults!

- *Compound-complex sentence*

  One or more main clauses containing one or more subordinate clauses; e.g., *I see Spot, who is my neighbor's dog, chasing a ball that belongs to Johnny.*

- *Complex-compound sentence*

  One or more main clauses containing two or more subordinate clauses that are coordinated; e.g., *I see that Spot is running and that he is chasing a ball.*

Section 4

Semantic Knowledge

semantics: *the nature, the structure, and the development and changes of the meanings of speech forms…* [38]

By this section, most of the structural foundation for using language, especially English, has already been presented. Linguistic knowledge about how to build words, phrases, clauses, and sentences provides a language user with the tools to express any idea. Merely possessing the tools, however, does not imply that a speaker can do everything possible with them. Indeed, this paper contains four more sections (including this one) on distinct types of knowledge that are also required to use language fully.

Although the concrete details are essential for language use, a *grammatically correct* sentence has surprisingly little to do with a *semantically meaningful* sentence, as Chomsky's classic example illustrates [27:10; 7]:

> *Colorless green ideas sleep furiously.*

Likewise, a semantically meaningful sentence may have little to do with grammatical correctness, as Yoda eloquently demonstrated in the movie *Star Wars, The Empire Strikes Back* [9:98]:[12]

> *When nine hundred years old you reach, look as good you will not.*

This section transitions from the concrete details of language building to the vague world of abstract meaning and its linguistic representation.

What is Semantics?

Semantics is the meaning of words and sentences. Specifically, within the framework of this paper, it refers to meaning that does not depend on the context where it appears.

[12] Yoda-speak is a rare (and of course, fabricated) example of an object-subject-verb (OSV) language, which probably gives it its exotic flavor. See *Clause Elements* on page 30 for more information.

Definitions of Semantics

Defining semantics is as difficult as working with it! Ultimately, the definitions lie in the eye of the beholder, as each field studying semantics has different goals, methods, applications, etc. Three views reasonably overlap the contents and purpose of this paper [23:371,373]:

- *Linguistics*

  Meaning is derived by syntactically analyzing surface structures of natural languages.

- *Philosophy*

  Meaning is considered in terms of logical propositions that characterize truth in arbitrary models and support certain methods of proof.

- *Computer science*

  Meaning is treated as the result of systematically executing commands in a programming language on a machine.

Views of Meaning

Several historical views of semantics have endured a long time. Indeed, the works of Aristotle, Plato, and Socrates are still referenced today for many purposes. Unfortunately, longevity does not necessarily imply correctness or usefulness in natural language processing. Three (flawed) conceptions of meaning form the backbone of further analysis [9:100; 51:192; 35:915]:

- *Naturalist view: word → things*

  In this popular view, originating with Plato, words supposedly name objects; e.g., *Las Cruces*, and *Chito*. For proper nouns, as well as certain closed semantic classes like colors, this approach works reasonably well. However, most words are simply too abstract to define this way; e.g., *to promise*, and *easy*, and *love*.

- *Conventionalist view: words → concepts → things*

  In this view, originating with Aristotle, words refer directly to objects, but they also have a conceptual reference, as shown in Figure 4.1. A variation on this view claims there is no direct reference, only the conceptual link. In either case, identifying "concepts" is simply too difficult for most words. Furthermore, even in cases where identification is possible, people may not share the same concept in mind; e.g., picture the many forms of a chair.[13]

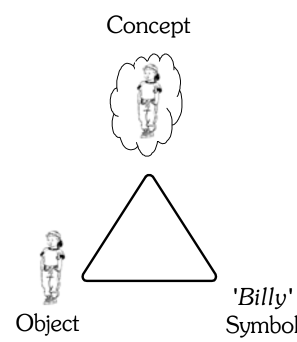


Figure 4.1: *Meaning Triangle*

[13] See [51:350] for a representative variety.

- *Behavioralist view: stimuli → words → responses*

  In this comparatively recent view[14], meaning is dependent on a chain of events: a stimulus causes someone to speak, and this speech triggers a response from someone else. While this process has some grounding in pragmatics and communication theory (e.g., *it's cold in here, isn't it?* could really mean *please close the window*), determining the dependent and independent factors in any response is intractable.

Modern linguistics tends to focus on how meaning is used within language as opposed to how it is linked with the external world [9:102]. This dichotomy alleviates some of the problems of the views above because semantic theories and semantic reality need not correspond in excruciating detail. The range of terminology and nuances vary significantly on this theme, but most overlap in two areas [9:102; 35:915; 27:216; 53:90,110,177; 54:16; 6:55; 26:381; 34:12; 58:9; 51:410,95; 23:392; 56:463; 2:256; 20:37]:

- *Sense or denotation or extension*

  The way the world really was, is, will be, or could be.

  Many aspects of the world have one unique definition in nature. For example, regardless of how a family tree is drawn or described, a person has only one biological mother and one biological father, and each recursively shares this property and so on, back to an arbitrary dawn of the human race.

  Other aspects of the world do not even exist. A unicorn, for example, is not real, but this technicality does nothing to prevent people from envisioning one and talking about it as if it were.

- *Reference or connotation or intension*

  The way people perceive or manipulate the world through language.

  Language parcels the world into conceptual chunks described by words. These words may overlap the physical world exactly, in which case sense and reference might be considered the same. More realistic[15], however, is that language provides a subset of reference words according to how it views the world. See Figure 8.1 on page 83 for a comprehensive example.

  Some languages may have specialized words for, say, *mother's father* (**mofa*), whereas others may have just the generic term *grandfather*. Both refer to the *same person*, but the latter also includes *father's father*. Neither reference can be considered more correct or incorrect, but they clearly have different connotations.[16]

[14] Leonard Bloomfield, 1933.

[15] Especially for reasons of practicality: having a unique word for every branch of a virtually infinite family tree, say *mother's father's mother's father*, would be utterly insane! More realistic is some arbitrary threshold where the branch loses its uniqueness through a generic term like *ancestor*. Social factors may play a role in defining this threshold, too; e.g., few people may care about details more than three generations back.

[16] Translating between languages that do not align their conceptual chunks in the same way is complex and fraught with difficulties [26:212,239; 11:332; 53:37,251,254; 54:74,132; 62:10; 23:46; 39:50; 45:162].

To complicate matters even further, many references are considered *vague* or *fuzzy* in that there is no clear-cut distinction between them [2:231; 23:402; 10:169; 20:10]:

How big does a watch have to be to become a clock?

How many grains of sand constitute a heap?

What distinguishes a hill from a mountain?[17]

When does a colt become a horse?

As a second example, consider all the little flavors, colors, spin, and so on that people put on references to make them fit the desired context. The following paraphrase about the CIA refers to the same group of combatants:[18]

If we support them, they're freedom fighters.

If we oppose them, they're terrorists.

If we're indifferent, they're warring factions.

Further discussion on reference can be found on page 88.

Deeper analysis is beyond the scope of this section. Consequently, the remainder of it is descriptive and practical in nature: it addresses key issues, observations, and problems about meaning without much regard to any particular theoretical or philosophical framework.

Sentence Meaning

The *Principle of Compositionality*, also known as *Frege's Principle*[19], forms the backbone for analyzing the semantics of sentences [27:224; 23:77]. In the earlier section on grammar (see page 28), it was shown how phrase structures combine to form sentence structures. The same principle applies to semantics: the semantics of phrases (roughly) combine to form the semantics of sentences [45:672].

Beyond the compositional meaning of its phrases, a sentence can also be analyzed on other semantic levels [9:107]:

- *Prosodic meaning*

  Meaning depends on the way a sentence is spoken. For instance, emphasis on certain words helps indicate focus or distinguish new information from old.

- *Grammatical meaning*

  The grammar of a sentence corresponds in many respects to certain meaning units. For example, the customary phrase order *Subject* + *Verb* + *Object* + *Adverbial* is often used in the sense *somebody* + *does* + *something* + *at some time*. *Thematic Structures* below investigates this connection further.

[17] Consider responses from people living on the east coast, in the Rocky Mountains, and in Nepal!

[18] Source unknown, vaguely reminiscent of something said in a Tom Clancy book or movie.

[19] Interestingly, Frege's works do not contain any explicit reference to this principle [56:420]!

- *Pragmatic meaning*

  Meaning depends on the context in which it appears. The next section on pragmatics covers this subject in detail.

- *Social meaning*

  The same relative meaning can be expressed with different social attitudes and implications; e.g., polite, reserved, rude, obnoxious, indifferent, partisan, etc.

- *Propositional meaning*

  Meaning can be translated into a logical form and manipulated in formal systems. This subject is investigated in Part II as a major component of knowledge representation.

Thematic Structures

Grammatical meaning, as described above, is often analyzed in terms the semantic roles shown in Table 4.1 [31:479,481; 2:248,251]. Linguists are not in complete agreement over how many roles exist, and other semantic labels are certainly possible. Sowa [51:508], for example, describes 19 roles, but many of them are just more finely grained distinctions; e.g., *origin*, *location*, and *destination* instead of simply *location*.

Role	Description
Agent	Volitional initiator of action
Patient	Object or individual undergoing action
Theme	Object or individual moved by action
Experiencer	Individual experiencing some event
Beneficiary	Object or individual that benefits from the event
Source	Object or individual from which something is moved by the event, or from which the event originates
Goal	Individual toward which the event is directed
Location	Place at which the event is situated
Instrument	Secondary cause of event; the object or individual that causes some event that in turn causes the event to take place

Table 4.1: *Basic thematic roles*

Ambiguity

Sentences are composed of phrase structures. The third property of phrase structures (from page 28), and therefore of sentences, is *ambiguity*: "Inherent ambiguity in language can be reflected in the grammar." Ambiguity, in its countless insidious forms, exists at every level of linguistics. Since a thorough survey of it would comprise an entire comprehensive exam of its own, this paper instead just touches upon representative examples.

No discussion would be complete without the classic example *time flies like an arrow*, which has at least the following interpretations [41:209]:

> *Time proceeds as quickly as an arrow proceeds* (the intended reading).
> *Measure the speed of flies in the same way that you measure the speed of an arrow.*
> *Measure the speed of flies in the same way that an arrow measures the speed of flies.*
> *Measure the speed of flies that resemble an arrow.*
> *Flies of a particular kind, time-flies, are fond of an arrow.*

Each of these interpretations is based on a different grammatical structure. The intended reading is the one most people would intuitively recognize as correct.[20] How people manage this task is an important practical question because computers must somehow match performance if they are to process natural language even remotely as well as humans do.

Word Meaning

As mentioned above, the meaning of a sentence is based on the meaning of its words. This discussion considers several important aspects of semantics at the word level.

Word Senses

Most words share more than one meaning or *sense*. In fact, for some heavily used words like the verb *take*, the number can exceed 80 [35:33]! Granted, many of the meanings are similar, but the fact that they are listed as separate senses indicates that some subtle distinctions exist [35:795; 2:231; 51:350]. Consequently, choosing the correct interpretation of each word is paramount to understanding the semantics of an entire sentence.

The formal process of correctly interpreting the contextual meaning of so-called *polysemous words* is known as *word-sense disambiguation*. Humans, for the most part, are not adversely affected by ambiguous words. Computers, on the other hand, have no special human abilities to handle them; as a result, poor semantic interpretation consistently degrades performance in natural language processing [11:332; 9:106].

Semantic Relatedness

One of the most studied aspects of semantics is how words are related to each other in meaning [27:220; 9:104].

[20] Indeed, few people would realize that so many valid (albeit less plausible) interpretations even exist.

Semantic relations

Most people are familiar with a variety of word relations, as they are commonly found in language resources like dictionaries or thesauruses [9:104, 27:220; 35:1015,483,73; 53:86]:[21]

- *Synonymy*

  The set of words that are relatively equivalent in meaning; e.g., *couch* and *sofa*, *start* and *begin*. Synonyms can usually substitute for each other, but since each has a slightly different connotation, one may be a more appropriate choice.

- *Homonymy*

  The set of words that are relatively equivalent in form in one of two says; e.g., *homophones* like *night/knight* or *flour/flower*, and *homographs* like *crane* (a bird) and *crane* (a lifting device). Such words lead to *lexical ambiguity*, which complicates the semantics of a sentence by providing multiple interpretations.

- *Antonymy*

  The set of words that are relatively opposite in meaning in one of three ways:

  – *gradable antonyms* are an open group of words with a range of scalar opposites; e.g., *hot* and *cold* versus *hot* and *cool*, *big* and *small* versus *big* and *tiny*. The degree can be qualified; e.g., *very hot*, *marginally cold*.

  – *nongradable* or *complementary antonyms* are a closed pair of words with polar opposites; e.g., *alive* and *dead*, *single* and *married*. The degree cannot be qualified in the literal sense; e.g., **partially dead*, **slightly pregnant*.

  – *converse* or *relational antonyms* are an open group of associated pairs of opposites; e.g., *parent/child*, *buy/sell*.

- *Incompatibility*

  The set of mutually exclusive opposites; e.g., *red* and *green*, in the sense that something cannot be both colors at the same time.[22]

- *Hyponymy or entailment*

  The set of words that are related taxonomically in meaning; i.e., X is a kind of Y, as in *dog* is a kind of *animal*. In logical terms, the following holds [27:221]:

  If something is in class A, then it must also be in class B (e.g., a *man* is a *human*)
  If something is not in class B, then it cannot be in class A (e.g., a *nonhuman* cannot be a *man*)

[21] Many other relations are possible when semantics is considered for knowledge representation. See [16] for an exhaustive discussion of WordNet.

[22] Of course, incompatible relations may have perfectly valid usages; e.g., something described as *red and green* generally has both *red parts* and *green parts*, which are not the same parts.

Semantic fields

Words can also be organized by their semantic associations or *fields*; e.g., an *insect* consists of a *head*, *thorax*, and *abdomen*, which respectively consist of *eyes*, *legs*, and possibly a *stinger*, etc. Roget's thesaurus laid the groundwork for such a taxonomy in 1852 by defining six top-level categories of words: *abstract relations*, *space*, *matter*, *intellect*, *volition*, and *affections* [9:104]. Figure 4.2 illustrates a small subset of the taxonomy, which extends to over 1,000 subcategories [9:104; 10:157; 35:914; 53:195].

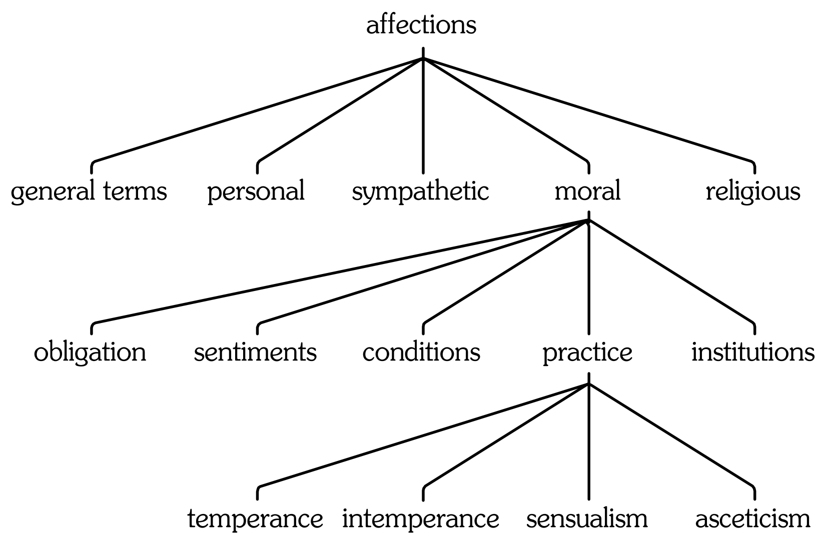


Figure 4.2: *Sample of Roget categories*

Semantic features

As the section on ontological representations of knowledge will discuss (see page 127), linguistic and taxonomic descriptions of the world are relatively arbitrary, and no single taxonomy can be considered "correct"[23] [45:232; 34:7; 51:409; 61]. The problem stems in large part from an incomplete understanding of the world and mismatches between it and language.[24] In other words, the semantics of many things is clear and straightforward, but others may fall into a gray area of uncertainty or a lexical gap in language [35:914].

[23] As the engineering saying goes, "all models are wrong, but some are useful" [51:384].

[24] Refer back to the discussion on referential meaning (page 37) for more information.

One way to represent how objects are related is with a feature matrix. For example, Table 4.2 indicates which English animal words correspond to which gender [9:107]. The last three animals are ambiguous in this respect.

	bull	*ram*	*boar*	*cow*	*ewe*	*sow*	*calf*	*lamb*	*piglet*
male	+	+	+	−	−	−	±	±	±
female	−	−	−	+	+	+	±	±	±

Table 4.2: *Sample feature matrix for animals*[25]

Likewise, Table 4.3 indicates which English verbs correspond to which manner of locomotion [9:107]. This example illustrates a lexical gap, as none of the verbs can directly represent *movement backwards*.

	walk	*march*	*run*	*limp*
natural	+	−	−	−
hurried	−	+	+	−
forward	+	+	+	+
one foot on ground	+	+	−	+

Table 4.3: *Sample feature matrix for verbs of locomotion*

Analysis of lexical features is especially important for translation. If features do not align between the source and target languages, then a relatively straightforward substitution of words with equivalent meaning is often not possible [11:332]. The meaning may have to be expressed in terms of different syntax, grammar, etc., and the human judgment calls and intuition involved in this process are daunting.

Qualifying Objects

As mentioned in the section on syntax (see page 17), nouns represent physical and conceptual objects in the world, and adjectives modify these objects. The semantic aspects of this linking are complex and defy straightforward rules and generalizations [27:224]. Consequently, this discussion considers only basic aspects.

Attribute meaning

If adjectives are properties and nouns are objects, and the principle of compositionality holds for both, then presumably the meaning of adjective–noun constructs should be some combined meaning of the parts. Indeed, this approach to semantics can be used to classify some adjective usage [27:224; 2:372; 45:708; 6:463; 51:81; 26:57]:

- *Pure intersective adjective*

  Many adjectives have a straightforward, absolute or "pure" meaning[26] that remains relatively the same regardless of the noun they modify; e.g., *blue*, *square*, *hollow*. For instance, a *blue*

[25] The markers + and − indicate the presence and absence of intersection, respectively.

[26] Although not necessarily unambiguous; e.g., *depressed*, *nerdy*, and *lacking personality* are alternative senses of these examples, respectively.

bird is roughly the same color as a *blue sky*, a *square hole* has the same number of sides as a *square peg*, and a *hollow sphere* and *hollow log* equally lack internal contents.

In mathematical terms, an adjective–noun construct is the set of everything with a certain property (i.e., adjective) intersected with the set of everything that is a certain object (i.e., noun). Figure 4.3 illustrates this operation applied to *green balls* [27:225; 34:66; 2:372].

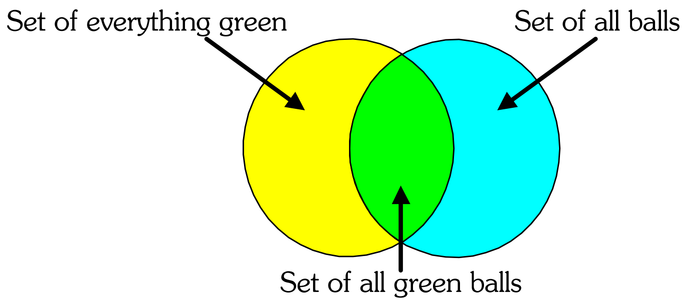


Figure 4.3: *Sample adjective intersection*

- *Relative intersective adjective*

  Many adjectives have a context-dependent or "relative" meaning that differs based on the noun they modify. For instance, a *short giraffe* rises far above a *tall human*; a *cold day* in Phoenix is a *hot day* in Fairbanks; and a *rich person* in Nigeria is likely a *poor person* in America.

  The same mechanism of intersection still applies as with pure intersective adjectives, but the context must somehow be factored in.

- *Uncertain*[27]*, nonintersective adjective*

  A special class of adjectives refers to conceptual descriptions that simply cannot be corralled into a set. For instance, an *alleged murderer* may not refer to a *murderer* at all. Likewise, a *possible solution* implies no guarantee that it is actually a solution. Context eventually leads to the proper interpretation, at which time such adjectives of uncertainty should no longer apply. For example, in **the alleged murderer was sentenced to death*, by legal definition the pronouncement of sentence removes any allegation, as the guilt is now considered fact.[28]

- *Nonreferential*[27]*, nonintersective adjective*

  A similarly special class of adjectives may not refer to actual descriptions at all. For instance, a *fake Picasso* is not even a *Picasso*; however, a *fake painting* is still a *painting*. The mathematical treatment of such sets appears untenable: the set of everything that is not really something intersected with objects that may not be themselves?!

- *Other adjectives*

  Some properties defy formal description; e.g., *an occasional cloud drifted by*. The meaning of this combination is clear—*a cloud occasionally drifted by*—but the usage does not conform

[27] My term, for lack of a better bulleted distinction.

[28] One would also hope that the murderer is actually guilty, but the veracity of the accusation has no bearing on the removal of *alleged* after sentencing (or the change to *convicted murderer*, as often happens).

to any of the categories above. In other oddball constructions, the adjective seemingly modifies the wrong noun. For instance, *a hot cup of tea* should actually read *a cup of hot tea*. Granted, an indirect result of hot tea within a cup is that the cup becomes hot, but one does not normally care about the temperature of the *cup* when drinking tea!

Attribute placement

If one adjective can modify a noun, then multiple adjectives obviously can, too. Moreover, one would not expect the meaning of the adjectives to differ based on their ordering. While this claim is generally true, also true is that certain orders are preferred, discouraged, or even prohibited [10:233]:

A nice, big, paper bag
* *A nice, paper, big bag*
* *A big, nice, paper bag*
* *A big, paper, nice bag*
* *A paper, nice, big bag*
* *A paper, big, nice bag*

No immediate generalization of these practically identical sentences is apparent, and upon further consideration, any attempt would probably not apply to other adjectives and nouns [10:233]. Such context-specific issues fall better into the next section on pragmatics. However, as they involve the semantic contribution of adjectives, they are discussed here.

Linguists have indeed been able to generalize some features of adjective ordering. While not without weaknesses and exceptions, the following microtheory suggests some predictability in four so-called *adjective zones* [10:223]:

$$\text{DETERMINER} + [\text{ADVERB}] + \text{ZONE}_1 + \text{ZONE}_2 + \text{ZONE}_3 + \text{ZONE}_4 + \text{NOUN}$$

- *Zone 1*

  Adjectives of absolute or intensifying meaning; e.g., *same*, *absolute*, *ultimate*, *extreme*.

- *Zone 2*

  Adjectives not falling into the constraints of other zones, basically everyday properties; e.g., *big*, *fat*, *slow*, *stupid*.

- *Zone 3*

  Adjectives of color or participial function; e.g., *red*, *blue*, *missing*, *stolen*, *destroyed*.

- *Zone 4*

  Adjectives collocated or associated closely with the noun; e.g., *wedding* [*dress*], *car* [*alarm*], *medical* [*facility*], *social* [*unrest*], *terrorist* [*attack*].

Certain noun compounds (see page 11) appear similar in meaning but function differently. Often, the role of the noun remains the same while the role of the property changes [35:245]:

> a *steamboat* is a boat that is propelled by steam
> a *riverboat* is a boat that travels a river
> a *gunboat* is a boat that has guns
> a *houseboat* is a boat that functions as a house more than a boat
> a *rowboat* is a boat that can be rowed

Similarly, the *Loch Ness Monster* lives in Loch Ness, whereas the *Cookie Monster* eats cookies. Such expressions are noncompositional and should be treated as single units of meaning.

Attribute scope

When multiples nouns are preceded by adjectives, the structure of the modification may become unclear [27:174]. This ambiguity is similar to that of noun-noun compounds (see page 11). For example, in *miniature badgers and raccoons*, are the raccoons miniature, or just the badgers? Disambiguating such lexical issues may require analysis on grammatical, semantic, and pragmatic levels. The details behind such analysis are beyond the scope of this paper.

Quantifying Objects

The various meanings of objects change in even more subtle ways when they are quantified by certain words. The influence of articles, in particular, is notoriously difficult to define.[29] In fact, the *Comprehensive Grammar of the English Language* devotes over 200 pages to *a* and *the* alone [43]! Quantifiers are considered here only simplistic terms [10:223]:

Definite article

- *Refers to the immediate situation*

  The object of reference is understood, perhaps because it is the focus of the context or is anticipated to become so; e.g., *the package has arrived*.

- *Refers to general knowledge*

  The object of reference is understood, perhaps because it is the obvious (or only) choice or it is mentioned often in similar contexts; e.g., *has anybody seen the dog*?

- *Expresses a specific state of affairs*

  The object of reference is unique within a class of objects; e.g., *Johnny wants to be the richest boy in the world* (meaning: of all the boys in the world, there can be only one richest).

- *Refers backward to an entity (anaphoric reference)*

  The object of reference was already mentioned; e.g., *John has both* [*a dog*]$_a$ *and a cat, but he plays only with* [*the dog*]$_a$.

[29] Even worse is trying to teach English article usage to speakers of languages that have *no* articles; e.g., Russian.

- *Refers forward to an entity (cataphoric reference)*

  The object of reference has not yet been mentioned; e.g., *John plays only with* [*the dog*]$_a$, *even though he has both* [*a dog*]$_a$ *and a cat*.

- *Refers to common human institutions*

  The object of reference is understood in general terms; e.g., *Bob went to the theater* focuses on his attending a movie, not on his going to a particular theater. Similarly, *Jack watched the news* refers to the notion of current events, not to specific events (unless so qualified; e.g., …*the news about the volcano*).

Indefinite article

- *Introduces an entity*

  The object of reference is unknown or uncertain in the immediate context, as illustrated by this narrative: *Jill says, "Look, there's an old lady outside." After peering through the window, Jack exclaims, "That's my MOM, you fool!"*

- *Expresses a general state of affairs*

  The object of reference is not being distinguished within a class of many such objects; e.g., *Johnny wants to be an astronaut*.

- *Often expresses a general quantity*

  The object of reference is a generic amount that is not a main focus in the context; cf. *the airplane used a gallon of fuel* versus *the airplane used the last gallon of fuel*.

No article

Many idiomatic constructions, especially those referring to common human institutions, appear without any article; e.g., *to go to bed*, *to have lunch*, or *at night*. However, they may allow the freedom to add such a distinction; e.g., *to go to the bed upstairs*, *to have a light lunch*, but not **at the night*.

Miscellaneous determiners

Words such as *all*, *every*, *most*, *some*, *few*, *no*, etc. can generally appear in the same contexts as the quantifiers discussed above. They impose semi-logical constraints on the objects they quantify; e.g., *all dog lovers and a few cat lovers but no more than three lizard lovers will be invited to the animal festival*. See *Types of pronouns* on page 16 for more information.

Special Meaning

Deixis and indexicals

A few special words play a role called *deixis* by referring to objects in ways that change somewhat predictably as the context changes [9:106; 12:20; 20:25]:[30]

- *Personal deixis*

  Pronouns refer to specific individuals or objects within the context of a conversation, etc. Once the context changes, the references usually change as well.

- *Spatial deixis*

  Certain words play the role of virtual pronouns by functioning on locations instead of individuals or objects; e.g., *come* <u>*here*</u>, *go over* <u>*there*</u>, *stand behind* <u>*that*</u>. As the context changes, the locations change, or the locations remain the same and the references change; e.g., *John tells Bill "Go over there to Jim." Jim yells to Bill, "Come over here to me."* Both *there* and *here* refer to the same place.

- *Temporal deixis*

  Other virtual pronouns function on time expressions; e.g., *yesterday*, *today*, *tomorrow*, *now*, *then*. As time passes, some—but not all—of these words refer to different times; e.g., *today* becomes *yesterday*, *tomorrow* becomes *today*, the *present* becomes the *past*, but the *past* remains the *past*, and the *future* may remain the *future* or become the *present*, etc.

Figurative language

Few aspects of language can ever be taken in the strict literal sense. Even worse is that sometimes communication is better served through *intentionally* nonliteral usage. Understanding such figurative language involves decoding the meaning(s) of its elements, as well as the relations binding them, in addition to the contexts of their normal separate usage and new combined usage. Pragmatic and world knowledge are essential for this task.

Many figurative constructions fall into just a few categories [10:172,177,421; 45:714; 35:653,739, 656,764]:

- *Metaphor or analogy*

  Different meanings of elements may be linked implicitly in terms of subtle association, comparison, and resemblance. For example, *the twilight of one's life* extends the notion *end of a day* to a lifetime, thereby meaning one's life is nearly over after its full duration.[31]

- *Simile*

  Metaphors may often be rewritten with explicit links, especially the comparisons *like* and *as*, for use in a more overt context. For example, *the boxer fought like a lion* transfers onto the

[30] Such context-sensitive, noncompositional meaning stands on slippery ground between semantics and pragmatics.
[31] With poetic overtones of impending darkness and the futility of trying to accomplish much at that point, etc.

boxer the property of how a lion would fight under similar circumstances. Indeed, for the sake of grammatical correctness at least, *the boxer fought as a lion would* is even more obvious.

- *Paradox*

  Elements that appear to contradict each other may be combined carefully to foster subtle insight into life; e.g., Charles Dickens's *A Tale of Two Cities* opens: "It was the best of times, it was the worst of times…"

- *Oxymoron*

  Elements that appear to clash opposite or incompatible notions may be combined for emphatic, humorous, or cynical usage; e.g., *bittersweet*, *jumbo shrimp*, *military intelligence*, *Microsoft operating system*.

- *Metonymy*

  An element may be replaced entirely by the name of something associated with it; e.g., *the White House announced…*, where the building substitutes for whoever actually made the announcement.

- *Personification*

  Properties or abilities of animate elements may be transferred abstractly to inanimate elements; e.g., *the gracious sun smiled on the hikers*, and *the rocks told the history of the canyon*.

- *Euphemism*

  Uncomfortable or inappropriate elements in many contexts may be replaced with milder forms; e.g., *to relieve oneself* in place of *to urinate*, and *adult video* instead of *porno movie*, etc. Political correctness and bureaucratic doublespeak have elevated many such expressions to the absurd; e.g., *aesthetically challenged* in place of *ugly*, and *unlawful or arbitrary deprivation of life* instead of *killing*.

Section 5

Pragmatic Knowledge

***pragmatics**: … meaning in context, or the meanings of sentences in terms of the speaker's intentions in using them* [38]

Morphology, syntax, grammar, and semantics all contribute heavily to meaning, and yet they appear to be little help in deciding the proper interpretation of this relatively simple statement [23:234]:

The osprey is looking for a perch.

The previous section on semantics briefly discussed word senses, or different meanings of the same word (see page 40). In many cases, context helps the reader determine the correct interpretation. Indeed, context does limit the choices here to only two:[32]

1 a. *The osprey is looking for a <u>fish</u>* [*to eat*].
  b. *The osprey is looking for a <u>branch</u>* [*to roost on*].

Both are correct interpretations. Which is the intended one? Nobody knows. Nobody can know because there is no context. If this sentence appeared in the context of, say, a story about *Flappy* the osprey, who has not eaten in days, then sentence (1a) would likely apply. Likewise, if *Flappy* has not slept in days, then sentence (1b) would. This analysis is considered further on page 69.

Incidentally, information alone is not the key to solving such ambiguity; rather, information *relevant to the context* is. For instance, knowing that *Flappy is gray with white, horizontal stripes and lives on an island inhabited by a lost race of untalented boy-band cavemen* ostensibly provides context for something, but not for determining the intended sense of *perch*.

Relevant context falls into the deep, dark, slippery realm of pragmatics.

[32] Only *perch* is of interest; any ambiguity of the other words is not considered.

What is Pragmatics?

Definitions of pragmatics vary as much as the subject material itself! The notion of context has arisen again and again throughout this paper as a way to generalize anything that does not fit the established framework. Such a definition is presented by many linguistic sources; e.g., "… anything relating to the way in which people communicate that cannot be captured by conventional linguistic analysis" [35:800]. Sowa [51:275] distilled a multitude of definitions into just two:

1. The basic meaning is a section of linguistic text or discourse that surrounds some word or phrase of interest.
2. The derived meaning is a nonlinguistic situation, environment, domain, setting, background, or milieu that includes some entity, subject, or topic of interest.

Context plays some role at every level of language. Correspondingly, many subdisciplines of linguistics overlap with pragmatics [10:120,116,22,38; 41:334]:

- *Pragmatics and Semantics*

  The intentions of a speaker and their effects on listeners are highly dependent on the context. The individual and common background knowledge, beliefs, presuppositions, etc. of the participants in communication carefully focus the usage of language.

- *Pragmatics and Stylistics*

  Language is infinite in the variety of constructions and meanings it can generate. Many common social contexts, however, are relatively limited in the forms that are considered "normal" or acceptable.

- *Pragmatics and Sociolinguistics*

  Language is a means of communicating states in an environment. Human social interaction comprises a huge part of this environment, so language and behavior have interdependencies. In many respects, sociolinguistics, stylistics, and pragmatics have a triangular relation.

- *Pragmatics and Psycholinguistics*

  Understanding how people generate and interpret language is of interest to linguists and psychologists alike. After all, the same brain performs all human functions, and its mechanisms probably overlap heavily.

- *Pragmatics and Discourse Analysis*

  Discourse, to be covered in the next section, forms the high levels of complex interaction between sentences. Not only do sentences have meaning alone, but they also form a larger picture together. Studies of literature, narrative structure, etc. intersect with pragmatics.

Communication

Giving and receiving information in some systematic manner is the basis of communication. At a minimum, two participants[33] are required to communicate. At least one must play the role of transmitter (e.g., speaker, writer, smoke-signaler), and at least one must play the role of receiver (e.g., hearer, reader, smoke-watcher). The minimum requirements establish *unidirectional* communication only; i.e., the transmitter sends messages to the receiver, who has no direct or immediate way to communicate back. Mass media—newspapers, television, etc., which are disseminated to a wide audience—generally follow this paradigm. Real-time conversation, on the other hand, is most effective if the participants play both roles to establish *bi-directional* communication, which allows direct and immediate response [23:94].

Communication Models

The way information is transferred is the subject of two popular models [45:658]:

- *Encoded-message model or semantic interpretation*

  The speaker has a proposition in mind and encodes it through language. The hearer receives the encoded proposition and decodes it. When this communication channel is free from error (e.g., noise or problems in encoding or decoding), what the speaker thinks and says is the same as what the hearer hears and understands.

  The meaning of the communication is the same each time it is transmitted.

- *Situated-language model or pragmatic interpretation*

  The meaning of a message depends both on the language used to encode it and the context or *situation* in which it is used. The human encoding and decoding "functions" are augmented with a context parameter.

  The meaning of the communication is different each it is transmitted with a different context. In fact, as it can be argued every context is unique to some degree, the meaning could be different even within the "same" context again.

Environment

The situated-language model tends to be more representative of the real world because it accounts for common contextual misunderstandings; e.g., *I thought you meant ____, so I did what seemed appropriate under the circumstances* [45:659]. Context is closely associated with two notions of *environment*, which is roughly the domain or world where communication occurs [23:75]:

- *Immediate environment*

  Objects of reference for the speaker and/or hearer are in the current environment; e.g., this week, every American knows at least some details about the destruction of the World Trade Center and can converse with anyone based on the shared context.

---

[33] Presumably, it could be argued that leaving oneself a note is also a form of communication, but such a case is slightly outside the scope of this section. However, it does indeed share many properties to be discussed; e.g., the context of such a note sometimes has a person wondering just what they were thinking when they wrote it!

- *Mediated environment*

  Objects of reference for the speaker and/or hearer are outside the current environment; e.g., a conversation about Abraham Lincoln refers to a context that no longer exists in the current world, only in the world of history. Nobody is still alive from that period, so all references are to some degree hearsay.

  Likewise, people born after the World Trade Center incident can refer to the immediate environment of their parents (e.g., *My dad said it was horrific and saddening*), but if they include only themselves, the reference may sound somehow odd; e.g., [*Although he was not alive to see it,*] *Billy said it was horrific and saddening* (how would he know?[34]).

Context

What roughly separates pragmatics from semantics is context, which plays some role in nearly everything imaginable. However, even with such an immense range and sphere of influence, most context falls into just a few categories [27:228]:

- *Physical context*

  Communication occurs in a certain place with associated objects and related actions.

- *Epistemic context*

  The participants in communication have certain background knowledge in common.

- *Linguistic context*

  Conversation proceeds as communication builds upon previous communication.

- *Social context*

  Participants in communication consciously and subconsciously behave according to social and cultural rules and norms.

Communication Actions

As mentioned above, the goal of communication is to transmit information. The range of this information is as infinite as are the grammatical structures that convey it; however, most information can be communicated as one of just a few discrete actions [45:652]:

- *Inform*

  Participants communicate unsolicited knowledge of the local environment to other participants.

- *Query*

  Participants ask other participants about their knowledge of the local environment.

[34] Much depends on the way the statement is made, of course. Anyone at any time can refer to an incident generically based on its timeless emotive value; e.g., *the destruction of Pompeii in* AD *79 was tragic*. However, alluding to direct personal emotional response (i.e., *psychological affect*) comes off as somehow peculiar; e.g., *I feel sad for the residents of Pompeii*.

- *Answer*
  Participants communicate solicited knowledge of the local environment to other participants.
- *Request*
  Participants ask other participants to do something.
- *Command*
  Participants tell other participants to do something.
- *Promise*
  Participants commit to doing something.
- *Offer*
  Participants negotiate deals with others to improve their own state or the state of the group.
- *Acknowledge*
  Participants respond affirmatively to requests and offers from other participants.
- *Share*
  Participants share knowledge and experience with other participants.

Communication Processing

The participants in communication usually do not express their ideas in a random, uncoordinated manner.[35] To do so would be very inefficient and prone to failure. The steps in communication actually conform to a relatively rigid framework that is repeated bi-directionally throughout a conversation [45:655].[36]

Speaker processing

The participant who starts the conversion or responds to the hearer assumes the role of speaker for the following encoding steps:

Step 1: Intention
The speaker wants the hearer to believe some proposition.

Step 2: Generation
The speaker generates the words that convey the meaning of the proposition.

Step 3: Synthesis
The speaker utters the words to the hearer.

Step 4: Role switching
The speaker becomes the hearer.

[35] *The Jerry Springer Show* and Fox's *The O'Reilly Factor* are notable exceptions!
[36] For a continuation of this discussion in terms of conversational discourse, see page 65.

Hearer processing

The participant who is not the speaker assumes the role of the hearer. He or she processes some of the information from the speaker as it arrives, while other information may be buffered for later processing or even discarded or lost. The following steps define the decoding process:

Step 1: Perception

The hearer perceives the words of the hearer, although not necessarily with perfect fidelity; i.e., details are frequently lost or mangled.

Step 2: Analysis

The hearer infers various meanings from the words and reconstructs the proposition.

Step 3: Disambiguation

The hearer infers that the speaker intended to convey the proposition (either the correct or misinterpreted version).

Step 4: Incorporation

The hearer decides to believe the proposition as he or she perceives it. Alternatively, the hearer rejects the proposition because it conflicts with his or her existing beliefs, assumptions, intentions, etc.

Step 5: Role switching

The hearer becomes the speaker. In conversations with more than two participants, any hearer can assume the speaker role. Extralinguistic factors like personality, assertiveness, domain expertise, personal experience, etc. play a role.

Speech Acts

People communicate in different ways to express different thoughts. These so-called *speech acts* reflect a deep understanding of explicit and implicit aspects of language, culture, social structure, etc. The many thousands of forms of expression can be distilled into just a few proposed categories [27:229; 9:121; 35:968; 11:353; 15; 19:26; 12:440; 20:64].

Direct speech acts

Stating a thought in an active, overt, or explicit manner tends to get directly to the point with little room for misinterpretation; e.g., *tell me what time it is* or *go to the store*. The recipient generally has few options in responding; i.e., to comply, to refuse, or to ask for more information before deciding.

Indirect speech acts

Stating a thought in a passive, subtle, or implicit manner tends to lessen the forcefulness of the statement. It also tends to make the focus less clear and more open for misinterpretation; e.g., *I wonder what time it is* or *Is milk on sale?* The variety of such constructions is bounded only by human creativity.

Other views on speech acts

Direct and indirect speech acts can be categorized differently according to their role in communication [35:624,968; 10:121; 27:229; 2:544]:

- *Locutionary speech act*

  The act of a speaker uttering a coherent string of words.

- *Illocutionary speech act*

  The act that is implied by a speaker in uttering a coherent string of words; e.g., *making a promise* implies future commitment.

- *Perlocutionary speech act*

  The act that actually occurs as a result of a speaker in uttering a coherent string of words; e.g., *telling a joke* should cause hearers to laugh, and yelling *Fire!* should start people running.

Finally, another categorization of speech acts focuses on the role of the speaker [35:968; 10:219; 2:557]:

- *Representatives*

  Speakers commit in certain ways to the truth of their utterances; e.g., *reporting*, *believing*.

- *Directives*

  Speakers try to convince hearers to do something; e.g., *suggesting*, *coaxing*, *urging*. A list is provided on page 32 for imperative verb forms.

- *Commissives*

  Speakers commit in certain ways to the consequences of their utterances; e.g., *promising*.

- *Declaratives*

  Speakers make announcements or proclamations that alter states in the world; e.g., *I hereby sentence you to 20 years hard labor*.

- *Expressives*

  Speakers express feelings or attitudes; e.g., *apologizing*, *congratulating*, *consoling*.

Felicity conditions

The effectiveness of speech acts, especially indirect ones, is often dependent on pragmatic requirements known as *felicity conditions* [9:121; 27:229; 6:227; 20:82]. In order for most people to comply with a request or command, etc., they must accept it as valid. For instance, anyone can state *I now pronounce you man and wife*, but only a certain set of people—justices of the peace, priests, ship captains—actually have the authority to back up such a statement.

Each type of speech act, as well as many verbs, has its own set of felicity conditions. For example, the following is a proposed skeleton for requests [15]:

- *Sincerity condition*

  When participant *A* requests that participant *B* perform an action *X*, the following must hold:
  - *A* must want *B* to do *X*.
  - *A* must assume that *B* can actually perform *X*.
  - *B* must be willing to perform *X*.
  - *A* must assume that *B* would not perform *X* without being requested to do so.

- *Reasonableness condition*

  Participant *A* must have a basis for making the assumptions of the sincerity condition. Unreasonable requests may be challenged on these grounds; e.g., *why the hell do you want me to do THAT, of all things?*

- *Appropriateness condition*

  The request must be appropriately framed with relevant information. A polite and friendly delivery with at least a perceived air of indebtedness helps as well.

Section 6

Discourse Knowledge

***discourse**: a connected series of utterances; a text* [1]

At the highest structural level of language, *discourse*, all the lower levels are assembled as a whole into coherent collections with related theme, meaning, or purpose. The form of these collections—both written and spoken—varies widely, from jokes to anecdotes, term papers to dissertations, short stories to epic sagas, books to tomes, and so on. What they all have in common, however, is some set of key ingredients, as well as rules and guidelines for combining them [10:286; 9:116].

What is Discourse?

Discourse connects individual thoughts into a big picture and communicates it in an appropriate manner. To this end, the speaker (or writer) has just a few major intentions in mind [45:718]:

- The speaker wants to convey a message
- The speaker has a motivation or goal in doing something
- The speaker wants to make it easy for the hearer to understand
- The speaker must link the new information to what the hearer already knows

Since discourse stands at the top of the compositional model of language, it demands an understanding of the knowledge contributing to all its lower levels [45:716]:

- General knowledge about the world
- General knowledge about the structure of coherent discourse
- General knowledge about syntax and semantics
- Specific knowledge about the situation being discussed
- Specific knowledge about the beliefs of the characters
- Specific knowledge about the beliefs of the speaker

Discourse Elements

Discourse appears in many forms with many goals; e.g., idle conversation, entertainment, teaching, preaching, historical analysis, political persuasion, and so on. Each form has its own unique features, which define both the structure in telling it and the expectations in hearing it. For example, in an action story, the main character usually gets into some dire predicament that materializes as the story progresses. In a good story, the listeners may not see it coming. Once it is apparent, however, the listeners expect some (exciting) resolution. The rare story that lets the main character die may be more realistic in real life, but it also tends to leave listeners disappointed since such a resolution conflicts with their expectations.

Storytelling is an age-old example of a relatively rigid discourse structure. All interesting[37] personal stories consist of six key ingredients known as *narrative macrostructures* [27:245; 9:119; 2:533]:

- *An abstract*

  A brief introduction to the story; e.g., *Man, my weekend was a complete disaster*.

- *Orientation or setting*

  A frame of reference giving the time, place, people involved, and so on; e.g., *Bob and I planned to go fishing up at the lake*.

- *Complication action*

  An unexpected element (or many) that complicates matters; e.g., *First we get a flat tire—and there's no spare! Then Bob gets a speeding ticket because his speedometer is broken. And—get this—when we finally get to the lake, we find it's been drained for seasonal maintenance!*

  This ingredient is similar to *conflict* in literature, where most stories fall into the categories of *man versus man*, *man versus himself*, *man versus society*, *man versus nature*, and *man versus the supernatural*.

- *Evaluation*

  The point of the story; e.g., *You can't believe our bad luck!*

- *Result or resolution*

  The conclusion; e.g., *So we went hiking instead. It probably turned out to be more fun than fishing anyway since I never catch anything but a cold on these trips.*

- *Coda*

  The moral of the story; e.g., *Next time, we'll get all the information—and a spare tire—before we go on a trip!*

[37] Uninteresting or pointless stories typically lack one or more ingredients, or they are told so poorly that the listener cannot determine the narrative structure.

Discourse Structure

Complex text cannot be represented as a linear ordering of independent sentences because together they would have little cohesive value [2:503; 11:5]. In other words, a collection of discrete sentences—even if they are all related to the same theme—is nothing but a collection of sentences; together they form no compositional structure with aggregate meaning.

Discourse is based on nonlinear or hierarchical ordering. Obvious examples are a table of contents and an outline, which specify overall structure at a high level. It is present at lower levels as well, which in turn may have even lower levels, and so on. Together they form the discourse structure, which is built from three components [45:717; 2:503]:

- *Discourse segment*

  A clause, sentence, or group of consecutive sentences forming a cohesive block of text in one of two ways [2:504]:

  – *intentional cohesion*, where sentences contribute to a common discourse theme

  – *informational cohesion*, where sentences are related by temporal, causal, or rhetorical relations

- *Coherence relations*

  Each discourse segment plays its own role and should be related somehow to the previous ones, future ones, and to the overall structure of the text. Relations vary widely, but most fall into just a few categories:

  – *causal relations* bind segments into a sequence of events leading up to some point.

  – *evaluation relations* connect segments but leave gaps in detail that must be inferred.

  – *elaboration relations* provide details to fill in otherwise inferred gaps between segments.

  – *explanation relations* provide background details or motivation for the current segment.

- *Enablement*

  Often a dependency or natural order of segments is imposed because, for instance, SEGMENT_2 cannot realistically occur until or unless SEGMENT_1 has occurred. In other words, one segment enables one or more other segments, possibly in a chain reaction.

Discourse Organizers

Discourse, like most aspects of language, is compositional. Control over the structure and flow of these elements is provided by discourse organizers [10:288; 2:518; 29:248; 28:133]:

- *Global macro-organizers*

  Discourse at a high level, such as topic and overall conversional structure, is directed by:

  – *topic markers* open a new topic; e.g., *What do think about President Bush?*

  – *topic shifters* move to a related topic; e.g., *On that note, what's the future of Alaska?*

  – *summarizers* close a topic; e.g., *Oh well, 2004 is just around the corner.*

- *Local macro-organizers*

  Discourse at a low level, such as supporting details and sentence structure, is directed by:

  – *exemplifiers* introduce supporting details; e.g., *for instance*, *for example*, *such as*.

  – *relators* connect statements; e.g., *moreover*, *furthermore*, *however*, *on the contrary*.

  – *evaluators* indicate position; e.g., *in my opinion*, *my view is*, *as far as I am concerned*.

  – *qualifiers* express uncertainty or non-commitment; e.g., *supposedly*, *from what I hear*.

  – *asides* provide control; e.g., *but that's beside the point*, *I'm getting ahead of myself*.

Grammatical connectivity

Syntax and grammar are generally considered to define only the structure of individual sentences. However, certain constructions also play a role in linking sentences in discourse [10:232; 9:119]:

- *Space and time adverbials*

  Elements linked in time and space; e.g., *The train left at sunrise. By lunchtime, it had already reached Chicago. When darkness fell, it was nearing its destination.*

- *Pronouns and other pro-forms*

  Elements linked through pronominal coreference (see page 15 and below); e.g., [*Jack and Jill*]$_a$ *went up the hill; they*$_a$ *were thirsty.*

- *Substitution*

  Elements linked through special coreference (see page 15); e.g., *Jack and Jill went up the hill for water*$_a$*; they found none*$_a$.

- *Determiners*

  Elements linked between general and specific objects (see page 46); e.g., [*A strange character*]$_a$ *wandered the park.* [*The man*]$_a$ *was later arrested.*

- *Comparison*

  Elements linked by comparing properties; e.g., *Three horses ran the race.* Speed Demon *was clearly the fastest* [*of the three*].

- *Connecting adverbials*

  Elements linked by sequencing; e.g., *Three issues must still be resolved. First of all, … Secondly, … And finally, …*

- *Conjunctions*

  Elements linked by coordination; e.g., *John liked the movie. However, Mary did not.*

- *Ellipsis*

  Elements linked by non-repeated elements; e.g., *John liked the movie. However, Mary did not*. Ellipsis is discussed below.

- *Repeated forms*

  Elements linked through explicit coreference, often for emphasis; e.g., *Bob Dole is running for President. Bob Dole wants your vote.*

Other connectivity

Other forms of connectivity transcend syntax and grammar [10:232; 9:206]:

- *General knowledge*

  Elements linked by relatively commonsense relations; e.g., *The weather was great this season. Farmers are expecting a bumper crop.*

- *Lexical relationships or vocabulary*

  Elements linked by related words; e.g., *That* $\underline{\textit{collie}}_a$ *is beautiful! She must be a nice* $\underline{\textit{dog}}_a$.

- *Punctuation*

  Elements linked by written features such as a comma, semicolon, colon, dash, etc.

- *Layout*

  Elements linked by graphical representations such as tables, charts, and diagrams, etc.

- *Prosody*

  Elements linked in vocal characteristics such as pitch, loudness, speed, rhythm, and pause.

Anaphora

Language and communication are generally efficient in conveying information without repeating unnecessary details. Pronouns are especially important for their ability to replace entire phrases, as the example on page 15 illustrated.

Pronouns can replace referents in two places [10:223]:

- *Anaphoric reference*

  The pronoun refers backward to an object that was already mentioned; e.g., [*The dog*]$_a$ *is hungry.* *He*$_a$ *has not eaten since this morning.*

- *Cataphoric reference*

  The pronoun refers forward to an object that has not yet been mentioned; e.g., *Because he*$_a$ *has not eaten since this morning,* [*the dog*]$_a$ *is hungry.*

Pronouns can operate at or beyond the sentence level [2:366; 11:112]:

- *Intrasentential reference*

  The pronoun refers to an object in the same sentence; e.g., *Bob*$_a$ *is an accountant, and he*$_a$ *really likes numbers.*

- *Intersentential reference*

  The pronoun refers to an object in a different sentence; e.g., *Bob*$_a$ *is an accountant. He*$_a$ *really likes numbers.*

  For the most part, only anaphoric references are used between sentences, but cataphoric references have their place, especially for introductions; e.g., *He*$_{a\rightarrow}$ *was the new villain in the old western town of Dirt Gulch. His*$_{a\rightarrow}$ *name was* [*Johnny Sixkiller*]$_a$. *And he*$_{a\leftarrow}$ *was one bad hombre, as the unsuspecting residents would soon learn.*

Pronoun usage is generally straightforward and predictable. For example, one study [25; 11:112] showed that between 90 and 95 percent of third-person singular pronouns—*he*, *she*, and *it*—refer to an antecedent in same or previous sentence. Furthermore, grammatical theories like C-commanding constrain pronominal interpretations [6:119; 2:367; 56:337].

Nevertheless, pronouns are a major problem for computers because semantics, pragmatics, and world knowledge in context have a monumental impact on interpretation [11:335]:

The monkey$_a$ *ate the banana because it*$_a$ *was hungry.*
The monkey ate the banana$_a$ *because it*$_a$ *was ripe.*
The monkey ate the banana because it$_?$ *was tea-time.*

Other pronoun-like constructions are equally troublesome [*ibid*]:

The soldiers shot at [*the women*]$_a$, *and* [*some of them*]$_a$ *fell.*
[*The soldiers*]$_a$ *shot at the women, and some of them missed*$_a$.

Certain features are helpful in decoding anaphoric expressions [11:112,119; 56:590]:

- *Syntactic information*

  Grammatical agreement, typically in gender and number, may constrain interpretations. For example, compare:

  John$_a$ *saw Mary*$_b$ *yesterday when he*$_a$/*she*$_b$ *drove by the store.*
  John$_a$ *saw Bill*$_b$ *yesterday when he*$_{\{a,b\}?}$ *drove by the store.*

- *Commonsense knowledge*

  Semantic agreement many constrain interpretations. For example, compare [60]:

  The organizers denied [*the protesters*]$_a$ *a permit because they*$_a$ *advocated violence.*
  [*The organizers*]$_a$ *denied the protesters a permit because they*$_a$ *feared violence.*

- *Salience*

  Context tends to continue, and correspondingly, the same referents remain the likely focus.[38]

[38] In careful writing, at least. Sometimes referents are linked indiscriminately; e.g., *The project was a disaster. It*$_?$ *was difficult to complete it*$_?$ *because it*$_?$ *was raining outside, and nobody knew what they*$_?$ *were doing.*

Ellipsis

Another form of anaphora is ellipsis, which omits otherwise repeated elements of a discourse structure. Resolving such omissions is based on two views [11:119]:

- *Syntactic view*

  A primarily syntactic element has been omitted; e.g., *John likes to fish, and Mary does* [*like to fish*], *too.* Similarly, *John is a lucky person. Me too*. (read: *I* [*am a lucky person*] *too*).

  Resolution involves inserting the omitted element in the proper syntactic form.

- *Semantic view*

  A primarily semantic element has been omitted. In simple terms, it is true that *John likes to fish*. If *Mary does, too*, then it is also true that she likes to fish. In contrast, if *Mary does not*, then it is false that she likes to fish (which may or may not be the same as saying it is true that she does not like to fish).

  Resolution involves inserting the semantic equivalent of the omitted element on a logical level that will be discussed in Part II.

Conversational Discourse

The discussion on pragmatic communication (see page 55) laid a framework for conversation in terms of the processing done by the speaker and hearer. This framework can be further described by the discourse mechanisms that operate on it [9:118]:

- *Opening*

  A plausible[39] beginning statement to a conversation; e.g., *nice car!* or *did you know?* or *I used to have one of those*.

- *Ongoing checks*

  A variety of cues, as shown in Table 6.1, that keep a conversation on track [10:289]:

  – by the speaker; e.g., *do you get it?* or *are you with me?* or the dreaded *y'know*?
  – by the listener; e.g., *uh huh* or *sure* or *by that you mean*… or *whoa, back up!*

[39] Pick-up lines are openings, but their plausibility is suspect! For example, *hey baby, what's your sign?* or *I lost my phone number; can I have yours?*

Expression	Example
Correcting oneself	*Actually, I meant Thursday, not Friday.*
Correcting someone else	*Don't you mean Thursday* or *isn't it actually Thursday?*
Requesting clarification	*Is this off the record?*
Requesting elaboration	*What's up with that?*
Recalling a forgotten question	*By the way, I was meaning to ask…*
Reorienting to new knowledge	*Sure, that's an option I didn't consider.*
Reacting to unanticipated information	*Oh, I didn't know that's even possible.*
Displaying recognition	*Uh huh, I've done that.*
Receiving new information	*You didn't think it would work, but guess what, it did!*
Marking intense reaction	*Oh yeah, you wanna bet on that?*

Table 6.1: *Sample conversional expressions*

- *Topic changes*
  - introducing a related shift in topic; e.g., *that reminds me…* or *incidentally…* or *speaking of that…*
  - indicating the end of a topic, but not the end of a conversation; e.g., *oh well!* or *that's life!* or *go figure!*
- *Ending*

  A preferably non-abrupt closing statement to a conversation; e.g., *well, it's getting late* or *I'd really like to talk about this more, but…*

Conversational maxims

Conversation is effective and efficient for many reasons. The basic conversational rules and guidelines within a culture, known as *conversational maxims*, are especially important because they establish a discourse framework [27:236; 9:117; 56:1187; 2:566; 20:88]:

- *Quality*
  - do not say anything that is believed to be false.
  - do not say anything that lacks adequate support .
- *Relation or relevance*
  - do say only something reasonably relevant to the topic.
- *Quantity*
  - do say something informative.
  - do not say anything too informative or overloaded with details.
- *Manner*
  - do avoid obscurity.
  - do avoid ambiguity.
  - do be brief.
  - do be orderly.

This list of maxims is not exhaustive, and new ones are easily added; e.g., *do be polite*, *do act consistently* [9:117]. Furthermore, many contexts and domains[40] have their own lists, which may even contradict this one [10:378].

[40] Politics is one such example, where opacity, obscuration, nonspecificity, etc. are the norm.

Section 7

World Knowledge

> ***world***: *all that concerns or all who belong to a specified class, time, domain, or sphere of activity* [1]

The section on pragmatics (see page 51) introduced an osprey looking for a perch. The two different senses of *perch* result in interpretations of the bird looking for either a *fish* to eat or a *branch* to roost on. Without further context, it is not possible to determine the *intended* interpretation. However, if this statement is viewed analytically, it may be possible to choose the *more plausible or likely* interpretation. Consider, for instance:

> How likely is it that a bird would be out looking for a particular species of fish? After all, if the bird is hungry, it will presumably eat any fish. Granted, it may prefer one species to another, but if hunger is the overriding factor, certainly any fish should do.[41]
>
> What is more important to immediate survival? All things being equal, food is certainly important, but impending exhaustion takes priority. The fact that the bird is flying at all means that it cannot be too near death from hunger. On the other hand, it could be close to falling out of the sky if it does not find a place to roost. To support this scenario, a context is envisioned where the bird is over water with no options to land.

Few people would formulate such detailed hypothetical scenarios to arrive at an interpretation. The point of this example is that people have the ability to consider the world in so many different ways. For all practical purposes, this analysis is not based on the definitions of anything in the original statement. Rather, it relies on an understanding of birds, flying, eating, resting, life, death, survival, priorities, preferences, and so on. In other words, it relies on knowledge of the world.

This final section is uncharacteristically weak and ill defined as compared to the previous six. World knowledge defies tidy definitions presented in neatly bulleted lists, etc. As a result, just an overview is provided.

[41] In this respect, birds may be more intelligent than humans: they probably would not do the equivalent of running out of a gas after driving all over town on an empty tank looking for the cheapest gas station!

What is World Knowledge?

For most English speakers, what comes to mind for the word "knowledge" is probably some subset of the listing in *Roget's Thesaurus* [36:290]:

ACQUANTAINCE

familiarity, awareness, understanding, apprehension, conversance, appreciation, consciousness, cognizance, realization, perception, enlightenment, experience, recognition, memory

EDUCATION

schooling, erudition, learning, scholarship, instruction, enlightenment

INFORMATION

facts, data, low-down, lore, science, wisdom

The *Oxford Dictionary and Thesaurus* augments this listing further [1:830]:

KNOWLEDGE

knowing, familiarity, awareness, apprehension, cognition, grasp, understanding, discernment, consciousness, conception, insight; ken, perception; facts, information, data, intelligence

An interesting observation is that none of these terms can be defined clearly and unambiguously. For example, how well must someone know something to claim *familiarity*, or how does someone become *wise*, or what is a satisfactory demonstration of *intelligence*? This problem is common to almost all definitions in language. It is further exacerbated by the fact that dictionaries circularly define words in terms of other equally vague and ambiguous words. The result is a form of "knowledge soup" that is based heavily on intangible aspects of human understanding [51:349; 45:320]:

- *Gut feelings*

  Some decisions just feel better than others. No empirical analysis can explain or justify them.

- *Generalizations*

  Language is effective and efficient because much of the obvious "stuff" is left unstated. Common sense and an understanding of the world fill in the gaps. For example, it is fair to state that birds fly. The fact that not all birds fly does not undermine this claim. Some birds—ostriches and penguins—do not fly by nature of their evolution; others birds—chicks, injured ones, and even dead ones—do not by nature of their current state. Nevertheless, they do not contradict the generalization, or default belief, about birds.

- *Abnormal conditions*

  By default, things that do something should be able to do it all the time unless unusual circumstances prevent it. For example, airplanes fly, unless the FAA grounds them, or an engine will not start, or bad weather makes it imprudent, and so on. There are few absolutes in the world.[42]

[42] Death and taxes are commonly cited exceptions, but only the former is a true absolute.

- *Conflicting defaults*

  The world rarely functions logically or according to definition; e.g., if Quakers are pacifists, and Republicans are not, then where did Richard Nixon, who claimed to be both, fit into this schema?[43]

- *Incomplete definitions*

  Some open-ended definitions have gaps that may not be worth filling or bridging; e.g., if a tree falls in a forest and nobody is there to hear it, did it make a sound? Often nobody cares about such hypothetical situations.

- *Unanticipated applications*

  Some open-ended definitions may lead to uncertain situations; e.g., if hair is part of the human body, then are hair implants, and wigs of natural hair, and so on? The dividing line is vague or undefined.

Knowledge types

A detailed discussion of knowledge will be presented in Part II (see page 83). However, for relative completeness within this section, two major distinctions are considered here [51:179,454; 23:494; 39:73]:

- *Semantic knowledge*

  General knowledge and beliefs about the world that are shared by most people regardless of the subject matter, domain, context, etc.

  For example, if John is *human*, then anybody—whether they know him or not—should assume that by default (i.e., unless it is demonstrated otherwise) he has countless human properties like *two arms*, *two legs*, *the ability to speak*, and so on. In other words, John is an ordinary, average human.

- *Episodic knowledge*

  Specific knowledge and beliefs about the world that are known only by people familiar with the particular subject matter, domain, context, etc.

  For example, say John's friends know that he was attacked by a shark years ago. One arm and both legs were eaten off, and he was left so traumatized that he no longer speaks. John is a specific instance of an ordinary, average human, but he is also different in some ways based on what happened to him. Perhaps only his friends know the whole story behind his condition; thus, they have unique episodic knowledge of what happened to him.

[43] Figure 12.3 on page 120 illustrates this conflict.

World Knowledge in Action

Most people do not give any thought to different interpretations in language. The fact that people rarely need to consider such details is a testament to its transparency. However, this ignorance can easily introduce problems based on different perspectives of the same situation. The following exchange describes a knowledge engineer trying to learn how a repair technician performs his job in order to develop a computer-based repair tool [51:355]:

> KNOWLEDGE ENGINEER: *I'm not sure what you want the system to do—determine whether a malfunction has occurred, determine what caused it, or determine what action to take in order to correct it.*
>
> REPAIR TECHNICIAN: *What's the difference?*

The technician likely deals with all three of these questions in his daily work, but he does not consider them separately. Experience, intuition, and other intangible aspects of knowledge direct his actions over the problem as a whole. As a result, it is difficult to provide a formal explanation for how or why people do certain things in certain ways. Even more difficult is to translate such vague, formless, tacit knowledge into the arcane formal structures required by computers.

A somewhat simpler example illustrates how people read between the lines to fill in implicit information:

1 a. *Sally got married and became pregnant.*
  b. *Sally became pregnant and got married.*

Sentence (1a) conforms to the culturally expected, ideal sequence of events. Sentence (1b), which is identical in outcome, implies something more than (1a) does; e.g., Sally may have felt compelled to get married as a result of becoming pregnant.

Finally, an extreme example that no computer will likely ever grasp is based on the following exchange [41:227]:

> WIFE: *I'm leaving you.*
> HUSBAND: *Who is he?*

The range of background information and subtle implications is immense! Deep knowledge is required about the institution of marriage, the reality of marriage, human nature, etc.

☐

Part II

Knowledge Representation

Part I of this paper focused on various aspects of linguistic knowledge. This part extends the discussion into how such knowledge can be formally represented and used by machines. Computers will never be human, but for them to process language as well as humans do, they must be embodied with many human qualities and abilities. The resulting system could be evaluated with the famous *Turing Test*, which challenges humans to determine whether they are communicating with another human or a machine. To pass this test, any so-called *intelligent agent* must be exhibit at minimum four capabilities (see also page 98) [45:5,157; 44:92; 23:52]:

- *Natural language processing*

  Natural language is the primary medium of communication, so mastery of it is essential.

- *Knowledge representation*

  Communication depends on general and specific knowledge of the world.

- *Automated reasoning*

  Knowledge is infinite, so no representation can store all of it. A means of inferring new knowledge must be available.

- *Machine learning*

  The world is dynamic, and new knowledge must be constantly acquired and adapted into the framework of existing knowledge.

Intelligent Systems

Research into intelligent agents varies according to the application. Most approaches work toward a goal of making systems that *think like humans*, *act like humans*, *think rationally*, and/or *act rationally*

[45:5,29]. These divisions reflect deep-rooted philosophical differences within the artificial-intelligence community, which finds itself firmly entrenched in two camps [45:29,818,827,839]:

- *Strong artificial intelligence*

  Intelligent agents can be conscious, sentient beings by virtue of their inherent complexity; i.e., brains make minds.

- *Weak artificial intelligence*

  Intelligent agents merely perform intelligently without any claims of consciousness or sentience. Such agents are simply good tools that get the job done.

Philosophical underpinnings and overtones notwithstanding, cognitive models to emulate or simulate intelligent behavior must exhibit many desirable properties by… [34:278; 44:228,477]

… being based on representations that actually represent
… adopting multiple approaches to representation
… using representations at multiple grain sizes
… being clear about the specification of processes
… attending to the details of processing as well as to its gross form
… attending to social context
… attending to relationship between the individual and the world

Representation Languages

Any agent—human or computer—that interacts intelligently with the world does so in a restricted manner because "all reasoning mechanisms must operate on representations of facts, rather than on the facts themselves" [45:158]. Therefore, the overriding issue is which form or language of representation to use [23:373; 51:178,419; 26:8,253; 12:14]:

- *Natural languages*

  Natural language is the ultimate language for knowledge representation. Everything that concretely, abstractly, or hypothetically exists in the past, present, and future can be richly expressed. To its credit, natural language… [26:8]

… is algorithmic with tractable computational complexity
… is general in its purpose and use
… combines verbal and sensory data
… makes effective use of logical contradiction and redundancy
… is inherently underspecified and context dependent
… facilities learning and knowledge acquisition
… allows reasoning over the environment and itself

On the other hand, natural language is also a poor formal language because… [12:14; 45:161]

… it cannot be easily manipulated algorithmically
… it is full of ambiguity
… its meaning is context dependent
… its syntax is extremely complex
… its connectives and prepositions are vague and unsystematic
… its supports few powerful rules of inference
… it represents spatial information poorly
… it evolved to meet the needs of communication, not representation

Surprisingly, several of these properties find themselves listed as both pros and cons. Iwańska and Shapiro [26:403] devote 11 pages to such issues.

- *Artificial languages*

  Artificial languages do not share the widespread understanding and usage of natural language, but they perform extremely well in certain highly formalized niches. Two types of languages belong in this family:

  - *programming languages* allow algorithmic execution of commands on machines.
  - *logical languages* prove the truth of propositions.

  Anything in natural language can be stated in programming languages and logical languages. The difficulty is in representing more than a small subset without running into overwhelming complexities.

Natural and artificial languages should not be considered competing formalisms; rather, they actually complement each other and play interdependent roles in many aspects of natural language processing. Two aspects are involved in mapping between formalisms [2:228]:

- *Semantic interpretation*

  A natural-language sentence is mapped to a logical representation, which straddles the boundary between natural and artificial languages.

- *Contextual interpretation*

  A logical representation is mapped to a knowledge-representation language, which is normally an artificial language.

The remainder of this paper focuses on artificial languages. To be of use in a system for logical reasoning, these languages must have four essential components [51:39; 44:947; 45:165]:

- *Vocabulary*

  A set symbols is needed to read and write expressions in a language:

  – *logical symbols* are domain-independent quantifiers, connectives, etc.; e.g., $\forall$ $\exists$ $\wedge$ $\vee$ $\Rightarrow$ $\Leftrightarrow$.

- *constants* are domain-dependent objects or individuals; e.g., *Bob*, *red*, *tall*.
- *variables* are symbols governed by the quantifiers; e.g., x, y.
- *punctuation* is symbols that clarify constructions; e.g., commas, parentheses.

- *Syntax*

  Well-formed expressions in a language must be built according to its syntax and grammar.

- *Semantics*

  The meaning of an expression depends on its components and their syntactic structure.

- *Rules of inference*

  A knowledge base is manipulated by applying rules to its facts. New rules are built from existing rules.

  Consistent with Ockham's razor—the most likely hypothesis is the simplest one that is consistent with all observations—logic systems of equivalent power can be compared according to the number of axioms they need to represent the same knowledge [45:534; 51:27]:

  - *weaker systems* use more axioms, so each must be less expressive.
  - *stronger systems* use fewer axioms, so each must be more expressive.

Logic systems appear in two types depending on how they deal with new information [12:101; 2:395; 55:668; 56:780; 51:373,381; 5:71; 45:326; 44:1022]:

- *Monotonic logic*

  Inferences are sound and stable, so everything a knowledge base entails can be proved without contradictions. Adding inference rules expands the knowledge base.

- *Non-monotonic logic*

  Inferences are made provisionally and may be retracted if overriding or contradicting facts are found. Adding inference rules either expands the knowledge base or causes existing rules to be removed. Updating the knowledge base is the responsibility of the truth maintenance system.

Section 8

Representation

Long before the Wright brothers ever flew a controllable powered aircraft, countless people envisioned flying machines modeled after birds. The rationale was simple: birds clearly fly, so flying is undoubtedly possible. Furthermore, birds fly by flapping their wings, so this approach should work equally well for machines. The results, of course, were laughable by today's standards. Nevertheless, this view of flight was a representation. It was inadequate in important respects (e.g., flapping is not the key) but correct in others (e.g., wings generate lift). Given enough time, effort, money, tools, etc. to theorize and experiment, early pioneers in aviation eventually refined their representation into a working form. As it was, the Wright brothers did not have a completely correct model of flight, either, so a representation need not be perfect to be useful—a representation simply needs to represent.

What is a Representation?

The world is full of objects that can be manipulated directly. For instance, a ball can be seen, touched, smelled, tasted, and heard. Granted, smelling and tasting a ball serve little purpose, and hearing a ball actually depends on its interaction with other objects, but all five human senses can indeed operate directly on a ball. Furthermore, experimentation is possible on a ball; e.g., throwing it at different angles has an observable effect on distance traveled.

A ball can also manipulated indirectly through a representation. A mathematical model[44] can view it in terms of radius, circumference, volume, and so on. In conjunction with the properties of mass, initial velocity, angle of departure, coefficient of friction, standard temperature pressure, and so on, distance traveled can be precisely calculated without ever throwing the ball.

[44] In engineering, geometric forms are wonderful models for approximation because oddly shaped real-world objects can be mathematically simplified. For example, the volume of a horse is much easier to compute if the horse is assumed to be a sphere! A more careful analysis would use multiple spheres. At some point, the spheres become small enough for an essentially exact solution. Volumetric knowledge representations are similar [12:271].

This example illustrates direct and indirect manipulation. For many objects, humans work only with the indirect representations or abstract models. For example, people cannot see electrons, yet electricity is well understood. Likewise, most people do not understand how an automobile works, yet they can operate it nonetheless.

Although their definitions, philosophies, implementations, and so on differ widely, all representations appear to share four components [34:5; 24:22]:

- *Represented world*

  A specific domain to represent must be chosen. Humans cover the gamut of domains, from the entire universe to the atomic level. Each has its own specific objects, properties, events, etc., along with tools for working with them and vernacular for talking about them.

- *Representing world*

  An artificial domain to parallel the represented world, or more often, a subset, must be chosen. For example, linguistics is a descriptive representation of language, but it does not account for everything in the domain of language.

- *Representing rules*

  A specific set of methods and rules to represent a domain must be chosen. Humans can use natural language for anything, but other methods may be more appropriate for some domains; e.g., mathematics and schematic diagrams.

- *Process that uses representation*

  A task or goal that uses the representation must be chosen. Representing something without using the representation is of limited value.

What is Knowledge Representation?

The short answer to this question is: anything that bridges the gap between the real world and its parallel models. Knowledge representations need not be complex. Bees, for instance, represent the location of food through a nifty little dance. This knowledge is *all* they can represent by it, but in the life of a bee, perhaps only food is important. Humans, with the most complex lives, have a mind-boggling array of representations to satisfy more advanced needs, as well as desires, beliefs, and other non-bee issues.

As knowledge representations are used in a wide range of tasks, they must exhibit a variety of useful properties [13:2; 136:134,121,15,1; 135:244,288; 135:165; 136:110; 133:48; 135:472,23; 134:90]:

- *Knowledge representation is surrogate*

  As described above, any manipulation of objects beyond physical reach must be done abstractly on something that stands in for them without actually being them.

- *Knowledge representation is set of commitments*

  Any domain of reasonable size or usefulness contains infinite detail. No representation can parallel them exhaustively. Therefore, decisions must be made about what goes into a knowledge representation and how much detail is appropriate. These decisions contribute to two intertwined commitments:

- *ontological commitment* considers the nature of reality and what exists in a domain. If only a subset can be represented, then presumably it should be a "good" subset. This value judgment depends on many factors within the domain and task, so no general-purpose answer exists.
- *epistemological commitment* considers the nature of knowledge and its justifications. The value judgments in ontological commitment directly impact what intelligent agents will be able to do with a representation. Committing to a subset may restrict its usefulness; e.g., if the comparisons *bigger* and *smaller* are not supported, then reasoning about size becomes difficult. However, if size is not important, then perhaps such comparisons can be safely omitted to reduce the size and complexity of a representation.

• *Knowledge representation is a fragmentary theory of intelligent reasoning*

Intelligent agents have many needs within a domain (see page 73). A representation must provide a means to satisfy them:

- *intelligent reasoning* constrains solutions to those that a human may reach.
- *sanctioned inferences* are solutions that pertain in some valid way to specific needs.
- *recommended inferences* are solutions that best pertain to specific needs.

• *Knowledge representation is a medium for efficient computation*

A representation that works "perfectly" in all respects but takes thousands of years to reach a solution is of theoretical interest, but it has no practical application.

• *Knowledge representation is a medium of human expression*

A representation should be relatively complete, understandable, and intuitive to everyone who uses it for whatever their purposes are. Knowledge engineers, domain experts, end users, and others should be able to communicate about it without undue difficulty. Certainly, the level of understanding and ability would differ among them, but nobody should be left out of the loop because they are not versed in the arcane "techno-speak" of the subject.

Levels of Representation

The remainder of this segment considers high-level issues in formulating and building a knowledge representation. It relies on the purposefully vague, non-technical term *stuff* to describe anything to be represented for any purpose.

Nature of stuff

The breadth and depth of stuff varies widely between domains and thus affects its representation [45:615]:

- *Explicit representation*

  Some knowledge lends itself well to exhaustive listing; e.g., a telephone book contains at least one phone number for each person. The names and most numbers have only an arbitrary correspondence, so there is no way to reason about them.[45]

- *Implicit representation*

  Most knowledge is so vast or loosely defined that it simply cannot be listed exhaustively. Instead, rules describe how it is accessed and manipulated:

  – *formulas* define explicit relationships over continuous or discrete intervals; e.g., $y=f(x)$.

  – *reasoning* provides a rational way to infer new knowledge from existing knowledge; e.g., dogs bark, and Fido is a dog; therefore, Fido barks.

Domain of stuff

The *environment* or *problem space* in which an intelligent agent operates depends on the task. Each type presents its own advantages and disadvantages in representation [45:46]:[46]

- *Accessible or inaccessible environment*

  An agent with complete information about an environment can generally make more informed decisions than an agent with partial information can. Too much information can be bad thing, however, because it may overload or overwhelm an agent.

- *Deterministic or nondeterministic environment*

  A deterministic decision is limited to a closed set of choices based on the current state and available information. Uncertainty is eliminated because all eventualities are known in advance.

- *Episodic or nonepisodic environment*

  Blocks of time in an episodic environment do not depend on previous blocks or affect subsequent blocks. Such independence is simpler to process than interdependent blocks.

- *Static or dynamic environment*

  An environment that remains the same while an agent deliberates over a decision is usually easier to deal with. In a dynamic environment, by the time a decision is made, it may no longer be appropriate or prudent.

- *Discrete or continuous environment*

  The steps toward a goal depend on the task. Discrete environments like chess execute in lock-step progression, whereas continuous environments like paintball run unfettered.

---

[45] Mnemonic correspondences like 1-800-CALL-ATT are a useful exception.

[46] Not surprisingly, the real world the most difficult environment; it is inaccessible, nondeterministic, nonepisodic, dynamic, and continuous [45:773].

Truth of stuff

Truth in representation is in the eye of the developer. For example, a perfectly implemented representation of a faulty theoretical model may work exactly as specified while having no valid connection with its counterpart in the real world. In other words, representations do what they are told, regardless of whether it is correct or not. In an ideally correct model, truth falls into two broad categories based on how the world really works [23:5; 45:821]:

- *External truth*

  The world exists regardless of whether any theories describe it. Long before humans came onto the scene, the same mechanisms were in play. These truths are constant; in a perfect world, science would devise exactly corresponding laws.

- *Internal truth*

  Human understanding of the world is incomplete; thus, so are the conceptual models that represent it. Occasionally, very convincing theories become laws; the remainder are continually re-evaluated, strengthened, weakened, or even discarded, etc.

  Two views subdivide internal truth:

  – *wide content* claims that a representation intrinsically refers to aspects of the real world by virtue of the fact that they can be represented at all.

  – *narrow content* claims no intrinsic connection; rather, the beliefs of the person who builds the representation play a role in specifying what it refers to.

Extent of stuff

The depth at which stuff is processed depends on how it is being used [26:283]:

- *Shallow representation*

  Some representations can operate well at a superficial level. For example, searching for names and dates in text can be done without analyzing syntax and semantics.

- *Deep representation*

  Many complex representations are so because they reflect inherent complexity in the world. Low-level processing must be employed to account for many details. For example, word-sense disambiguation must often consider both syntax and semantics.

- *Mixed representation*

  Both shallow and deep representations have their places, and they are not mutually exclusive. In most cases, processing should be done only at the level needed for a particular task, etc.

Form of stuff

All representations operate on some form of symbols instead of on actual stuff. These symbols can appear in two fundamentally different forms [34:9,12; 51:285; 44:90; 11:677,807;786; 53:88; 51:69]

- *Symbolic representation*

  Computers and most programming languages function on the symbolic level. Variables assume a fixed set of values and are manipulated through functions with predefined, known behavior. For example, digital circuits operate on discrete signals through well-defined logic functions. In binary logic, the expression ($a \wedge b$) maps four inputs to two outputs; no other states are possible. Developing, testing, and tweaking such circuits involves a finite (although often large) number of cases. The compositional nature of discrete logic allows modularized architectures to be tested as relatively independent units.

  Symbols are often recursively composed of lower-level symbols down to atomic *primitives*. The *symbol-grounding problem* addresses difficulties in defining which primitives exist and what they mean. A dictionary exhibits an analogous problem: entries are defined in terms of other entries, so somebody who understands none of them cannot use it.

- *Subsymbolic representation*

  So-called *soft computing* is valuable in many representations where imprecision and uncertainty are common and making discrete approximations carries a significant penalty. Artificial neural networks and fuzzy logic (see pages 129 and 110, respectively) are common implementations.

  Disadvantages of subsymbolic representations are manifested in development and maintenance because the internal mechanisms are hidden and not well understood. For example, analog circuits operate on continuous signals through loosely defined logic functions. In fuzzy logic, the expression ($a \wedge b$) maps an infinite set of inputs to outputs. Even in discretized increments of 0.1, 100 inputs map to 10 outputs. Developing, testing, and tweaking such circuits involves an intractable number of cases.

Implementation of stuff

The amount of detail in stuff may ranges from superficial to gory specifics. Ultimately, however, the implementation level needs to have its requirements satisfied, so low-level details must be produced at some point. This abstraction can be considered on multiple levels [45:153,257; 34:21; 13:7]:

- *Knowledge or epistemological level*

  What does an agent know? For example, if it understands decimal numbers, can it add?

- *Logical level*

  How does an agent encode what it knows? For example, to add, does it use prefix, infix, or postfix notation in an imperative or functional statement?

- *Implementation level*

  How does an agent actually do what it knows how to do? Three lower-level abstractions apply:
  - *computational-level descriptions* can be treated as black-box functionality, so internal details are abstracted away; e.g., an adder unit takes two inputs and outputs a sum.
  - *algorithmic-level descriptions* considers intermediate details with mid-level abstraction; e.g., an adder converts the decimal inputs into binary numbers, performs an addition instruction, then converts the result to the decimal output.
  - *implementation-level descriptions* considers all details; e.g., the addition instruction is performed by a 32-bit carry-lookahead adder built from discrete logic.

Represented Knowledge

Knowledge appears in so many forms that even a survey is well beyond the scope of this paper. Instead, this segment extends the discussion on world knowledge (see page 71) by considering just two high-level categories.

Semantic Knowledge

General knowledge and beliefs about the world are known and shared by most people regardless of the subject matter, domain, context, etc. Such knowledge is not innate to humans, as it must be learned at some point in life. Nor is it the same for every human, as people learn different things in different ways, and not everything is even correct [44:947]. Nevertheless, semantic memory—the storehouse of this semantic knowledge—reflects the world in a way that can be considered reasonably common and static [51:179,454; 23:494; 39:73; 12:436,453; 136:110]. For example, people familiar at all with liquids recognize that they appear in various forms with different properties, as illustrated in Figure 8.1 [51:353; 12:342]. While most people cannot exhaustively list all these forms, and in fact, not all languages even make such distinctions (see page 36), a tacit understanding of liquids is arguably shared.

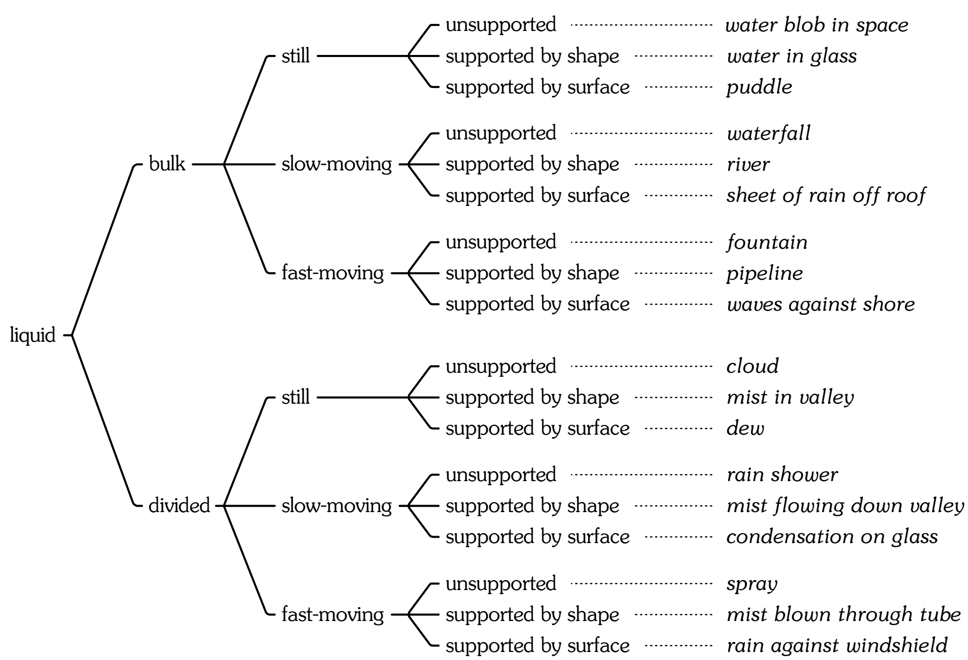


Figure 8.1: *Forms of liquids*

Semantic memory is based on at least five components [136:454]:

- *Ontology*

  The world consists of interrelated generic concepts with properties. The discussion on ontology (see page 127) provides more details.

- *Definitions*

  Formal rules and conditions dictate how concepts are organized within an ontology and how new ones are described in terms of existing ones.

- *Constraints*

  General principles or axioms specify what must be true of everything in an ontology.

- *Defaults*

  An ontology supports the so-called "rules" in the real world, as well as the exceptions to them.

- *Behavior*

  Among the properties of a concept are what it can and cannot do, what can and cannot be done to it, and so on.

Episodic Knowledge

Specific knowledge and beliefs about the world are known only to people familiar with the subject matter, domain, context, and so on. The range of people can vary from an individual to the entire human race; the defining factor is knowledge of particular objects, places, events, etc. Whereas semantic memory maintains relatively constant shared knowledge such as *dogs are related to wolves*, episodic memory maintains transient knowledge about, say, my dog *Rambo* and the wolf *White Fang* from a Jack London novel.

Episodic knowledge revolves around *episodes*, or instances and events in the world [51:454]. As each instance is unique, a knowledge representation must account for this variation. Furthermore, changes over time must be considered, as very little in the world is constant. These issues are tackled by the theory of *fluents*, which leads to formal representations like *temporal logic* and *situation calculus* (see page 111) [45:241; 12:191,58; 56:917].

An event or instance that occurs once is unique; if it occurs more than once, each is similar yet different in some respect. Two elements of representation can account for these differences [12:191]:

- *State types*

  Many statements of fact are timeless in the sense that they answer the generic question: *was something ever true at any time, including the present*? For example, Godzilla has clearly been in Tokyo: *in*(*Godzilla*, *Tokyo*).[47]

[47] A good knowledge representation should distinguish between reality and fantasy. As no such creature actually exists, and the guy-in-a-lizard-suit instance that is referenced was actually in a mock-up of Tokyo, this "fact" is valid only within the context of certain movies. To compound the problem, references to fantasy are commonly used in the real world; e.g., *If the Japanese don't lower their trade barriers, we're going to send Godzilla after them!*

- *State tokens*

  In many cases, a timeless representation is inadequate because multiple instances of the same event must be distinguished. For example, Godzilla actually frequents Tokyo, and to refer to a particular attack requires each to have a unique name of some sort. For example, the tokens *token*(*attack*$_1$, *in*(*Godzilla*, *Tokyo*)) and *token*(*attack*$_2$, *in*(*Godzilla*, *Tokyo*)), as well as the relation *before*(*attack*$_1$, *attack*$_2$), can be used to define specific instances like *during*(*attack*$_1$, *crushed*(*Godzilla*, *tanks*)) and *after*(*attack*$_2$, *went*(*Godzilla*, *volcano-lair*)).

The order of events depends on when they occur relative to each other. Since not all contexts require the same level of detail in representing this order, different dependencies are useful [51:215]:

- *Implicit dependency*

  In some contexts, it may be appropriate to define events in the vague terms of *always* or *sometimes*. If something is always true, then it applies to every fact; however, if something is true only sometimes, then it cannot be expected to be true before, during, or after any particular event. In order words, *always* means *it is necessary*, whereas *sometimes* equally means *it is possible*. See the discussion on higher-order logic (page 109) for more details.

- *Explicit dependency*

  In other contexts, exact points or intervals in time may be needed. Typed variables can be defined on a timeline and compared relatively or absolutely; e.g., $\forall t_1$:*time* $\exists t_2$:*time* *during*(t_1,t_2) $\wedge$ *occurred*(e,t_2) means some event e occurred at some point t_2 within the (unspecified) interval t_1. Temporal relations are shown in Figure 9.5 on page 101.

Events may form many widely differing kinds of chains [51:216,115,125]:

- *Discrete or continuous events*

  Computers operate in discrete lockstep with a fixed-interval clock that unambiguously defines exactly when things can occur; e.g., the minimum quantum of time may be a clock cycle on the order of one nanosecond. The real world imposes no such limit on granularity; i.e., for any moment in time, a smaller unit of time can be found. By convention, humans view the world in terms of seconds, minutes, hours, and so on, but these units are arbitrary and infinitely divisible along a continuous interval.

- *Linear or branching events*

  A chain in the past is linear because only one branch was chosen at each alternative point. For future chains, however, all branches are possible, and the number of potential outcomes explodes combinatorially.

- *Independent or ramified events*

  Independent events have no effect on other events, whereas ramified events do in two ways:

  - *locally ramified* events effect only those things with which they are directly involved.
  - *structurally ramified* events cascade effects across things regardless of their involvement.

- *Immediate or delayed events*

  Some events cause an immediate effect, whereas others serve as a trigger for consequences later down the chain.

- *Sequential or concurrent events*

  Some events must occur in linear order by time, whereas others may occur simultaneously.

- *Predicable or surprising events*

  If all the effects of an event can be listed exhaustively, then its role in a chain is predictable. Effects that are not known in advance or not considered introduce uncertainty.

- *Normal or equinormal events*

  Many chains play out in an expected or more probable order, whereas others may be entirely unpredictable.

- *Flat or hierarchical events*

  Events in a flat chain are dependent on a limited number of preceding events. In a hierarchical chain, events are composed of (possibly recursive) subevents that must propagate their effects forward in time.

- *Timeless or time-bound events*

  Timeless objects in a chain exist within its entire context (and possibly beyond), whereas time-bound objects may wink in and out of existence as conditions change.

- *Memory-bound or forgetful events*

  The next event in a chain may depend on the preceding events, or it may be independent of everything except the current event.

Definitions of Knowledge

Commitments must be made about many issues in a knowledge representation before anything can be represented. While this paper focuses on formalized approaches of arcane representations, knowledge need not appear at this level. In fact, many high-level resources like a dictionary, thesaurus, and encyclopedia are used by ordinary people. They are indeed representations of knowledge, and, as such, they should also conform to certain commitments and guidelines [58:16]:

- *A definition must be neither too broad nor too narrow*

  An overly broad definition includes objects that are not part of the term being defined; e.g., *a dog* is *a hairy quadrupedal animal* applies equally well to wolves. Likewise, an overly narrow one excludes objects that conform to it; e.g., *mammals* are *animals that give birth to live young* does not account for the duck-billed platypus. A definition can be both too broad and too narrow.

- *A definition should not be negative when an affirmative definition is possible*

  A description of what something is not fails to capture the properties of what it is; e.g., *an equilateral triangle* is not *scalene or isosceles*.

- *A definition should be literal*

  Figurative language (see page 48) is too vague; e.g., *diamonds* are *a girl's best friend*.

- *A definition should be evaluatively appropriate*

  Emotive or biased language like sarcasm is inappropriate; e.g., *Microsoft* is *a company that produces bug-free software; yeah right!*

- *A definition must be mutually non-circular*

  Recursive chains contribute little or no information; e.g., *a prize* is *something awarded after a contest*, and *a contest* is *something that awards prizes*.

- *A definition must be consistent with previous definitions*

  No contradictory or anomalous cases should arise from substituting equivalent definitions; e.g., if *a pencil* is *made of wood* and *an automatic pencil* is *made of plastic*, then a plastic automatic pencil must also be made of wood!

- *A definition must be unambiguous, not excessively vague, and not obscure*

  Ambiguous terms should be avoided, or they should appear with adequate context; e.g., *Wells Fargo* is *a bank (a financial institution)*. Vague and obscure language clouds the meaning; e.g., *kittens* are *good animals*.

Knowledge can be defined in many arbitrary ways. One common way is by the intended purpose or use of a definition [58:10]:

- *Stipulative definition*

  Some definitions simply connect or repackage other definitions; e.g., a ≡ b means a *is equivalent to* b; or, *a spork* is *a combined spoon and fork*.

- *Lexical definition*

  Language-related resources like a general-purpose dictionary define words in terms of their conventional or customary meanings; e.g., *a spoon* is *a small, shallow, oval-shaped bowl on a handle used for stirring and eating food*.

- *Precising definition*

  For specific purposes, especially legal contexts, vague definitions must be explicit; e.g., *intoxicated* means *drunk*, but legally, it means *a blood-alcohol level of at least 0.8*. Note that someone with 0.79 may be drunk in the general sense while not in the legal sense.

- *Theoretical definition*

  For a particular audience, a definition may be limited to the context or framework of a theory; e.g., for a physics student, *cold* is *reduced kinetic activity at the molecular level*.

- *Persuasive definition*

  Words with connotative or emotive baggage (see page 37) can bias a definition; e.g., *the members of the northern group* are *freedom fighters*.

Another way to define knowledge is by its representation [58:12; 51:92,99,583] (see also page 37):

- *Extensional or denotative definitions*

  The meaning of an extensional definition is based on the objects to which it applies:

  – *enumerative definitions* list every member of the defined set; e.g., *the U.S. states* are *Alabama*, *Alaska*, *Arizona*, and so on. Many sets contain too many members to list or the members are unknown, etc.

  – *ostensive definitions* refer to actual objects; e.g., *Chito* is *<u>that</u> person* (indicated by pointing). This approach is limited to physical contexts where all elements are present.

  – *recursive or inductive definitions* rely on a proven starting point and rules for proceeding based on it; for example:

  someone's parent is someone's ancestor:

  $\forall x\ \exists y\ parent\text{-}of(y,x) \Rightarrow ancestor\text{-}of(y,x)$

  the parent of someone's parent is someone's ancestor:

  $\forall x\ \exists y\ \exists z\ (parent\text{-}of(x,y) \land ancestor\text{-}of(y,z)) \Rightarrow ancestor\text{-}of(x,z)$

  nothing else is someone's ancestor:

  $\varnothing$

- *Intensional or connotative definitions*

  The meaning of an intensional definition is based on the properties that its corresponding objects must have:

  – *synonymous definitions* indicate that multiple definitions are generally equivalent; e.g., *a physician* is *a doctor*, and vice versa.

  – *contextual definitions* are synonymous definitions that appear more appropriate in certain contexts; e.g., $(a \land b)$ is the same as $\neg(a \lor b)$.

  – *operational definitions* list steps toward proving that an object exhibits the required properties; e.g., *a liquid* is *an acid* if (and only if) a clean litmus strip placed in contact with it turns red.

  – *hierarchical definitions* use the notion of *genus* and *species* in relation to other definitions. A genus specifies objects with like properties, and a species specifies how these objects are different; e.g., the genus *Canis* contains wolves, dogs, foxes, and jackals. Each is a separate species because it differs in notable characters.

References to Knowledge

References connect elements of the world to the various types of symbols that represent them [51:396]:

- *Elements applied to signs:*
  - *qualisign (material quality)*
    beeping is a function of sound waves, regardless of their source or meaning.
  - *sinsign (material indexicality)*
    beeping from the direction of an alarm clock indicates the probable source.
  - *legisign (material mediation)*
    the purpose of an alarm clock is to wake sleeping people, so its beeping sets a context.
- *Signs applied to objects:*
  - *icon (relational quality)*
    the image of an alarm clock evokes understanding and meaning of its real-world counterpart.
  - *index (relational indexicality)*
    an arrow pointing toward the image of an alarm clock brings it into focus within a context.
  - *symbol (relational mediation)*
    the sound of alarm clock in a different context may be used in related contexts.
- *Objects applied in language:*
  - *rheme (formal quality)*
    the arbitrary English word *alarm clock* refers to such an object, whether it is present or not.
  - *dicent sign (formal indexicality)*
    reference to a specific alarm clock establishes an actual context.
  - *argument (formal mediation)*
    a sequence of dicent signs establishes a process involving an object.

Reference in natural language is in many ways analogous to pointers in programming languages. Not surprisingly, many of the same problems manifest themselves. For instance, two people can refer to the same object by a different name and not realize it. Likewise, they can refer to different objects by the same name. Two main issues arise in representing reference [2:380; 23:391; 45:244; 12:55; 56:463]:

- *Transparent reference*

  Equivalent terms can be substituted freely within the appropriate scope. For example, given *Kazi is the whitest kitty*, and *Kazi bit Dan*, it is safe to assume that the whitest kitty bit Dan. Both descriptions refer to the same kitty. Leibnitz's principle formalizes this as "… two expressions refer to the same object if in all contexts they can be interchanged without changing the truth value" [56:463].

- *Opaque reference*

  Equivalent terms cannot be substituted freely within the appropriate scope. For example, given *Kazi is the whitest kitty*, and *Iris believes that Kazi bit Dan*, it is not safe to assume that Iris believes the whitest kitty bit Dan because Iris may not know that Kazi is the whitest kitty. Iris has an explicitly stated belief only about the biting. In fact, she may know of a kitty that is even whiter than Kazi, in which case the substitution would have that particular kitty biting Dan.

Section 9

Reasoning

Reasoning is the process of drawing conclusions through the plausible manipulation of knowledge and information. Knowledge representation and reasoning are integral parts of each other, but they can be addressed separately in terms of what they are and how they is used [135:92]. This section discusses the latter in various kinds of reasoning [134:25,95; 135:140; 51:247,143]:

- *Representing degrees of belief*

  People know how well they know many things; e.g., experts are likely positive about their answers, whereas novices may hedge theirs with a degree of uncertainty.

- *Evaluating the strength of arguments*

  People usually believe things they know or infer until something indicates that they are incorrect or makes them less certain.

- *Applying rules of general but not universal validity*

  Statements about the world are generally true, but exceptions are nearly always possible. The presence of an exception does not undermine the validity of generalizations or belief in them.

- *Avoiding the enumeration of all conditions on a rule*

  An infinite number of remotely plausible factors are relevant to any situation, but sane people almost never consider them. For example, a person who wants to buy a soft drink from a vending machine operates under the basic assumptions that everything will go according to the way it normally does. While it is indeed possible that a power failure could occur or the money could be counterfeit or a runaway truck loaded with mutant clams could crash through the wall, planning for such eventualities is pointless. On the other hand, certain eventualities are so common that contingency planning is almost obligatory; e.g., vending machines consistently rip people off, so taking extra coins may be prudent.

- *Inferring from the absence of information*

  The default interpretation of many unknown situations is usually the broadest, most general possible; e.g., hearing an unknown dog bark would lead someone to believe that a dog is nearby, but not that the dog is white, belongs to a rich person, and has rabies.

  In other situations, the default interpretation is assumed false because, as Carl Sagan stated in debunking the supernatural, "extraordinary claims demand extraordinary proof." If a person does not know the correct answer and maybe even has no clue whatsoever, he or she may still reject a claim on the grounds that it conflicts with common sense, conventional wisdom, or personal beliefs.

  Finally, information often provides its own plausible constraints somehow; e.g., if concert tickets are announced to cost at least $50, people would not normally infer that they cost $1,000, even though this conclusion is mathematically true.

- *Limiting the extent of inference*

  Plausible rules on a small scale may lead to implausible results on a large scale. For example, it seems intuitive to claim that removing a grain of sand from a heap leaves behind a heap. However, after repeated applications of this rule, the single grain of sand remaining is decidedly not a heap.

- *Inferring with vague concepts*

  The heap example above illustrates the ill-defined semantics of vague concepts and the resulting difficulty in reasoning over them. The discussion on reference (see page 37) provides more examples.

- *Finding expected utility*

  Life involves continuously re-evaluating the cost of doing something with respect to its potential benefits and the possible consequences of not doing it. For example, carrying an umbrella in case of rain incurs the cost of finding it, holding it, not losing it, etc. The benefit is enjoyed only in the event of rain, whereas the consequence of not carrying it is getting wet.

- *Reasoning about reasoning*

  Reasoning about the pros and cons of a decision, as in the umbrella example above, demonstrates a form of abstract, second-stage thinking. This meta-reasoning reveals to rational agents some of the ways they make their decisions and control their fate.

- *Inferring explanation*

  Everything happens for a reason, but the reason is not always apparent. People fill in the gaps based on available information and commonsense reasoning. For example, finding one's car wet when the sky is overcast generally leads to the conclusion that it rained. Seeing kids with a hose next to the car on a sunny day generates another conclusion. The combination of both an overcast sky and hose-wielding kids requires more information and thought.

- *Inferring schemas*

  Many aspects of life and the world exhibit regularity and predictability. Inference on at least four related components helps fill in missing or anticipated information:

  - *schema identification* involves determining the context of a situation; e.g., seeing a fuselage, wings, and tail leads to the conclusion that an object is an airplane.
  - *slot prediction* involves inferring the expected components of an identified schema; e.g., seeing an airplane leads to the conclusion that it has a cockpit and seats inside.
  - *filler identification* involves putting actual instances into slots; e.g., Captain Crunch sits in the cockpit and named passengers occupy the cabin.
  - *relation prediction* involves predicting what the parts of the schema likely do; e.g., the engines somehow make the plane go because engines do so in general.

- *Inferring by analogy*

  Knowledge of something in one context or schema often applies to others. For example, engines make cars go, so, by analogy, they should also make jet airplanes go. The fact that two entirely different mechanisms provide the power for two different ways of going does not undermine the usefulness of this analysis.

- *Inferring a general rule from examples*

  People naturally jump to unsound conclusions based on minimal information or experience; e.g., Bob thinks all the food at Lo Wang's Chinese restaurant is bad because he did not like the one tiny bite of kung fu shrimp he tried.

Inference

Reasoning involves selectively applying inference rules that lead to other rules and eventually to a desired conclusion, if one can be drawn [55:672]. Reasoning in the real world must take into account countless vague interdependencies between extremely complex and ill-defined elements, etc. While such reasoning is undoubtedly possible—human do it easily—it demands far more varied and powerful mechanisms than are normally available in today's knowledge representations. As a result, examples of reasoning usually apply relatively simple inference rules.

Most inference rules are based on three approaches [12:4; 2:393,470; 12:4; 51:390; 5:129,141]:

- *Deductive Inference*[48]

  Beginning with something true and following true intermediate steps to a conclusion is proceeding deductively. Only conclusions that follow a valid line of reasoning can be proved; e.g., rocks can break glass, so throwing a rock at a window could break it because it is glass. Saying rocks cannot break glass causes the remainder of the chain to make no sense, as does throwing a rock at a tree, etc.

[48] See Figures 9.2 and 9.3 on page 100 for more information.

- *Abductive Inference*

  Finding plausible reasons for why something is true works backwards toward the source. Multiple interpretations may be equally valid; e.g., finding one's wallet gone could mean it was stolen or lost or misplaced, etc. Each line of reasoning develops differently.

- *Inductive Inference*[48]

  Encountering multiple examples of something naturally leads to generalizations; e.g., if a person in a new town sees several purple fire hydrants, it does not take long to generalize that all fire hydrants in town are purple. This reasoning does little to develop a hypothesis about why they might be purple, but it does help predict the color of future fire hydrants to be encountered. Furthermore, after a generalization has been made, a person is less likely to consider the issue much further; e.g., the first purple fire hydrant was probably a total surprise, and to a lesser degree, so were the following couple; after a certain small number, however, the surprise is gone.

Theorem Proving

Applying inference rules in a strategic way to reach a goal forms the basis of *theorem proving*. Many variations in the details are possible, but most are based in some way on the basic methods discussed here.

Forward and Backward Chaining

Chaining relies on the fact that the inference rules are tied to other inference rules through logical implication. Rules in the chain $p \Rightarrow q$, $q \Rightarrow r$, $r \Rightarrow s$, $s \Rightarrow t$ and $t \Rightarrow u$, for instance, can be proved by two chaining methods based on the available information and intended goal [44:687; 34:93; 45:274; 51:156]:

- *Forward chaining*

  Following a chain from start to end relies on the rule *Modus ponens* and positive logic: if x is known to be true, then by applying $x \Rightarrow y$, infer that y is true. Therefore, in this example, knowing premise p is true means that conclusion u is true as well.[49]

  This inference is not directed toward any particular goal; rather, it proves everything that can be inferred from the premise, even irrelevant facts.

  Most people intuitively use forward chaining in commonsense reasoning.

- *Backward chaining*

  Following a chain from end to start relies on the rule *Modus tollens* and negative logic: if y is known to be false, then by applying $x \Rightarrow y$, infer that x is false. Therefore, in this example, knowing conclusion u is false means that premise p is false as well.[49]

  This inference is directed toward a particular goal by proving everything from it on a directly related chain back to the premise. It returns all relevant answers to a question.

  Few people use backward chaining in commonsense reasoning.

[49] As well as q, r, s, and t, of course, but as intermediate steps, they may not be of interest.

Chaining is used to explain existing conditions and to predict future ones [45:447]:

- *Diagnostic inferences*

  Working backward from effects to presumable causes is the basis of diagnosis; e.g., what would cause a patient to have green skin and swollen earlobes?

- *Causal inferences*

  Working forward from causes to possible effects plays a role in prediction; e.g., what might be the consequences of doing *x*?

- *Intercausal inferences*

  Multiple causes can play a role in an effect, but with limitations; e.g., two independent causes are less likely to occur simultaneously than one is, so their combined role should play a much smaller role in determining potential consequences.

- *Mixed inferences*

  Any combination of diagnostic, causal, and intercausal inferences is possible; e.g., an ill person involved in a car accident might have multiple unrelated medical conditions with common, disjoint, or conflicting symptoms, etc.

Entailment and Implication

Natural language often involves reasoning over properties and limits in at least two ways [57:56; 2:254,393; 26:272]:

- *Entailment*

  A more restrictive statement can be inferred as true from a less restrictive true statement. For example, *all dogs bark loudly* entails *all dogs bark* because to bark loudly requires them to bark in the first place. Likewise, *few birds speak* entails *few birds speak English* because if few speak at all, then presumably even fewer speak a specific language. This statement is true even if all the few birds do indeed speak English.

- *Implication*

  A statement can be constrained for practical interpretation without enforcing the constraint. For example, *Larry has two dogs* leads to the natural conclusion that he does not have three dogs, but he could.

Other Methods

The following methods warrant mention as they are commonly employed in inference and reasoning. However, detailed discussion would delve into proofs and other aspects beyond the scope of this paper.

- *Resolution*

  Probably the most used inference rule in artificial intelligence, resolution builds new rules—usually too many—by mixing and matching other rules in a pattern-matching manner; for example [55:658]:

ball(*marble*)	a marble is a ball.
$\neg$*ball*(*x*) $\vee$ *round*(*x*)	a marble ball cannot not be a ball, so discard the offending term, which results in a round marble ball.
$\neg$*round*(*x*) $\vee$ *rolls*(*x*)	a round marble ball cannot not be round, so discard the offending term, which results in a round marble ball that rolls.
$\therefore$ *rolls*(*marble*)	therefore, the marble rolls

- *Unification*

  The pattern-matching of resolution can be extended to evaluate functions and compare lists of arguments to compute the most general solution based on a given line of reasoning [4:270; 2:411,599,604,607].

- *Circumscription*

  Properties that apply to the majority of objects can be made the default by abductively minimizing the set of nonconforming objects; e.g., describing *most birds fly* as *the set of non-flying birds is as small as possible* [56:789; 12:109; 2:395; 5:141].

- *Subsumption*

  Redundant rules can usually be collapsed into more general ones to reduce their number in a knowledge base; e.g., *person*(*x*) subsumes *person*(*Bob*) because, say, if everyone is a person, then there is no point in stating that Bob is a person [45:286].

Reasoning

To function in the real world, higher organisms must reason in some way over many aspects of life. Humans are the most rational and intelligent of all organisms and correspondingly have the most complex lives. This segment considers several of the main ways humans reason.

Commonsense Reasoning

Much of what makes humans human is the ability to use so-called *common sense*. The gap between human and computer abilities in reasoning is vast, and it is unlikely to be spanned until a better understanding is formed about what common sense is and how it works [12:18].

Some properties of commonsense reasoning contribute to its flexibility and tolerance to varying situations [44:687; 56:590]:

- *Reasoning with incomplete info*

  Common sense typically applies whenever it is needed, even if not all relevant information is available. Humans excel at filling in the gaps, entertaining multiple hypotheses, etc. based on what is available, expected to be available, and even certain never to be available; e.g., aircraft accident reports often must draw conclusions from missing, damaged, or destroyed flight data.

- *Reasoning with changing info*

  Humans excel at reformulating reasoning on the fly as further information becomes available. New or supporting information builds a line of reasoning, whereas contradicting information may change it; e.g., a free Sunday afternoon offers multiple possibilities for outdoor leisure depending on the weather. If the weather changes after a plan has been adopted, it may be modified or even discarded in favor of, say, something indoors.

- *Reasoning with uncertainty*

  Real-world information can rarely be taken at face value. Reasoning must account for differing validity, strength, salience, and so on to develop or maintain plausible lines while discarding others. For example, planning on Monday for a Sunday trip leaves open many possible weather-related factors. Contingency plans can be devised in advance and reformulated as the weather on Sunday becomes more predictable.

Logical reasoning systems tend to fail on real-world situations because the real world does not lend itself well to formal description. The fact that humans do not fully understand the world contributes to poor representations of it [45:417]:

- *Laziness*

  Too much manual work is required to consider all possible cases within a real-world scenario. As a result, many details are consciously omitted in building a reasoning model.

- *Theoretical ignorance*

  Many real-world scenarios are simply not understood well enough to consider all the possible cases. Moreover, even if they are recognized, formally representing them may be beyond the current state of the art.

- *Practical ignorance*

  An exhaustively defined reasoning model may work beautifully on complete, valid, constant data known in advance, but, of course, the real world rarely conforms to these requirements.

Reasoning about Beliefs

Lying, cheating, selling used cars, and other self-serving human endeavors rely on intentionally misrepresenting the world to an agent who is unaware of the deception [44:950]. Belief, or the epistemic state, of agents represents their understanding of the world at a particular time based on available information and reasoning [5:48]. Some beliefs are shared within a scenario; e.g., both a used-car salesperson and a potential customer understand that a car is to be purchased. Other beliefs are unilateral;

e.g., the salesperson believes the customer will pay more than necessary for a car, while the customer (occasionally) thinks the salesperson will be honest [2:548]. The customer may entertain multiple beliefs simultaneously; e.g., used-car salespeople try to rip customers off, but not all of them do, so perhaps give this one in the checkered sports coat and striped slacks the benefit of the doubt. Some beliefs may be perceived (maybe incorrectly) as more likely than others; e.g., most used-car salespeople are crooks, so make this belief the default expectation [34:72; 12:96,370].

Belief revision is the strengthening, weakening, or discarding of beliefs as more knowledge is acquired; e.g., a rational agent can eliminate any doubt once the salesperson pitches "Have I got a deal just for you, Roger! Don't tell my manager, but I'm gonna sell you this beauty for less than cost!" [44:1241,1244] Linguistic mechanisms behind beliefs are discussed in the segment on speech acts (see page 56).

People have an unending supply of false beliefs that may or may not be revised throughout life. Their form varies widely and their logical soundness is often questionable. True beliefs in a sound reasoning system are more limited [2:546; 12:357; 20:133]:

- *Explicit beliefs*

  Belief propositions are explicitly stated as true in the knowledge base; e.g., Chito is human.

- *Implicit beliefs*

  Belief propositions can be inferred from explicit beliefs in the knowledge base through inference rules; e.g., Chito is a man, and a man is human; therefore, Chito is human.

Beliefs play a major role in human life. In fact, BDI (belief, desire, intentional) models are considered essential to building any truly intelligent agent [2:542]:

- *Perception*

  An agent must receive and process information about the world around it.

- *Beliefs*

  Information on the present state of the world must be represented.

- *Desires and wants*

  Certain states should be preferred to others so that the desirability of beliefs can be compared.

- *Planning and reasoning*

  Changing states depends on understanding past, present, and potential future states, as well as their cascading effects on other agents and the world.

- *Commitment*

  The decision to change states should be made rationally.

- *Intentions*

  Decisions in a chain toward a goal should remain relevant.

- *Acting*

  An agent should be in active control of these actions.

Case-Based Reasoning

Reasoning in law provides the foundation for *case-based* or *analogical reasoning*, which compares known cases and their conclusions to new cases to draw similar conclusions. Figure 9.1 depicts the basic architecture of a case-based reasoning system [34:225]. The *case base* contains known cases in indexed form for quick and flexible lookup by the *retriever*. The *tweaker* massages the new case to fit into the framework of the known ones. The *learning/storage* unit processes the cases to determine whether their conclusions can be reconciled. It outputs the result and records it as a new case in the case base.

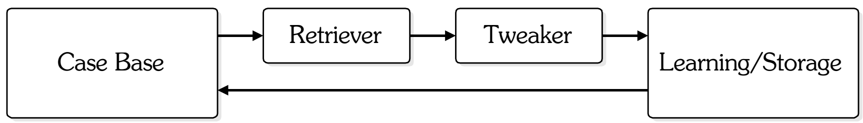


Figure 9.1: *Case-based reasoning architecture*

Case-based reasoning derives conclusions through four components [51:358; 44:196]:

- *Given case*
  A known case with a conclusion is used as a precedent; e.g., smoking crack cocaine is illegal.
- *New case*
  A new case is found to be similar to the given case; e.g., is smoking the new smackhead rot cocaine illegal?[50]
- *Cause*
  An attempt is made to reason over the new case in light of the given one by ironing out their differences; e.g., smoking crack makes people high and dangerous to society; smoking smackhead rot cocaine leads to the same result.
- *Judgment*
  If the differences between cases can be reconciled satisfactorily, some part of the original conclusion can be drawn for the new case; i.e., smoking smackhead rot cocaine is illegal.

Case-based reasoning can record both successful and unsuccessful results [45:829; 34:201]:

- *Success-driven learning*
  Conclusions that work for a given cases are stored for future use on the same or related cases. Generating them again will not be necessary.
- *Failure-driven learning*
  Conclusions that do not work for a given case are stored for future use as negative evidence. By knowing which cases did not lead to a solution, fruitless reasoning can be avoided.

[50] The ideal of precedent extends only so far toward the legal realm. Drugs, for example, are described in law by their chemical formula, which need only be changed slightly to make it a "legal" variant of an illegal drug.

Figures 9.2, 9.3, and 9.4 compare how some of the reasoning schemes discussed in this section learn from cases [51:359].

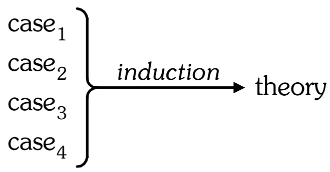


Figure 9.2: *Induction*

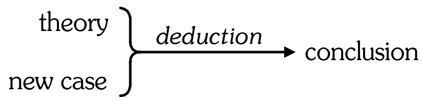


Figure 9.3: *Deduction*

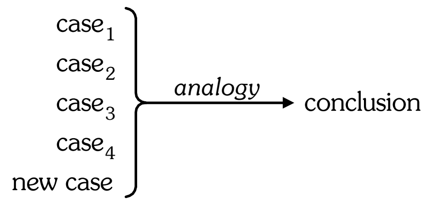


Figure 9.4: *Analogy*

Temporal Reasoning

Time moves inexorably forward, and with it moves everything in the world. Reasoning about the world thus requires a clear understanding of time and the ways that language represents it. Furthermore, any temporal reasoning must operate on complex, interrelated times and durations in the actual and hypothetical past, present, and future [12:187].

Although time is constant[51] throughout the universe, its representations vary [48:31; 2:410; 51:115; 12:149]:

- *Metrical time*

  Time is represented explicitly as numerical instants along a timeline. Standard mathematical relations and operations hold over the intervals; e.g., t_2 is earlier than t_5 by three time units.

- *Topological time*

  Time is represented implicitly in terms of relative relations like those shown in Table 9.5 [3].

[51] Einstein's Special Theory of Relativity notwithstanding.

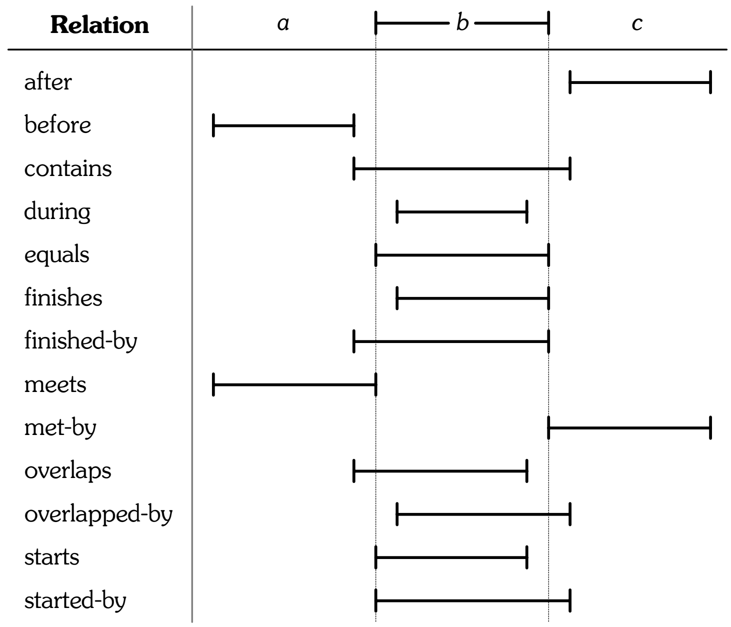


Figure 9.5: *Temporal relations*

Temporal relations combine in various ways to define ongoing processes, completed events, and static states [51:214]:

- *Continuous process*
  Changes occur incrementally without periods of inactivity.
- *Discrete process*
  Changes occur in indivisible steps.
- *Continuous-process initiation*
  Every process has a starting point.
- *Continuous-process continuation*
  Every process has some intermediate points between its starting and ending points.
- *Continuous-process cessation*
  Every process has an ending point.
- *Discrete-process event*
  The indivisible steps from initiation through cessation comprise an event.
- *Discrete-process state*
  Periods of inactivity during events are static intervals where nothing changes.

Figure 9.6 illustrates the relationships between these elements [51:214]. Periods of activity and inactivity appear as squiggly and straight lines, respectively.

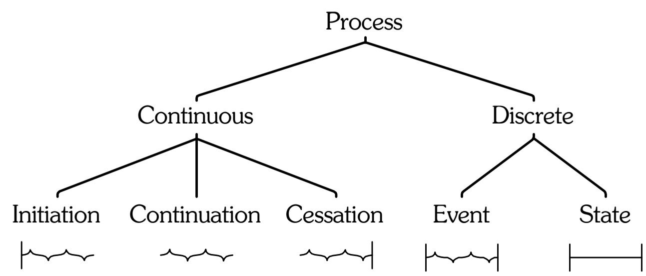


Figure 9.6: *Processes*

Processes can also be considered in terms of their fundamental nature [48:13]:

- *Random process*

  In total chaos, nothing is predictable, and the random states in any "process" have no relationship to each other.

- *Law-governed process*

  Much of the world is predictable at some abstract level. For example, the law of gravitation helps to reason that a handful of marbles will fall when they are dropped, but it does not predict where each will hit the ground.

- *Deterministic process*

  Certain scenarios are completely predictable according to tightly constrained laws. For example, opposite poles of a magnet always attract and like poles always repel.

Points and intervals along a timeline represent different temporal propositions [2:407; 56:903]:

- *Stative proposition*

  States that do not change over an interval or any subinterval are static.

- *Activity proposition*

  On-going actions may change at any time.

- *Telic proposition*

  Completed actions do not change any further. They can reach a final state in two ways:

  – *achievements* focus on the transition between states.

  – *accomplishments* focus on the activities leading up to the final state.

Spatial Reasoning

Natural language makes heavy use of common prepositions like *on*, *in*, *by*, *near* and so on to build spatial representations of the real and imaginary world.[52] These relations, along with spatially deictic words like *here* and *there* (see page 48), are especially troublesome to reason over because they depend on unseen context [135:950; 135:2; 134:40,251,273]. Representations of moving objects may involve temporal reasoning as well.

Although vision is not considered part of natural language processing, the two share several properties [134:241]:

- *High-level vision*

  Interpreting images involves understanding what is being seen in the given context.

- *Physical reasoning*

  Reasoning is performed over the shapes, positions, and movement of objects based on what they are and what they can do, etc.

- *Route planning*

  Determining how a moving object might change location involves understanding where it is going, why, how, etc.

[52] Many of the same prepositions play a role in temporal relations, too.

Section 10

Mechanisms of Reasoning

Declarative Approaches

This section, as well as the next two, provides a survey of popular formal mechanisms to represent knowledge. Each is arranged by implementation:

- *Declarative approaches*

  Representation focuses on the objects and their properties and interrelations within a model. Inferences are made after defining what the objects are and what they can do.

  This symbolic approach is based on the declarative and functional programming paradigms of languages like Prolog and LISP.

- *Procedural approaches*

  Representation focuses on ways to assemble and manipulate objects algorithmically.

  This symbolic approach is based on the imperative programming paradigm of languages like Java and C.

- *Graph-based approaches*

  Representation focuses on graphical models with hybrid utilization of both declarative and procedural approaches.

  Computer science defines no corresponding programming paradigm; instead, these symbolic and subsymbolic approaches borrow from engineering, mathematics, graph theory, psychology, philosophy, cognitive science, and other disciplines.

Propositional Logic

Many declarative sentences in natural language can be treated as logical expressions since they state facts, (which need not even be true); e.g., *Chito is a Ph.D. student*. In systematic combination with other "facts," conclusions can be drawn; e.g., *Ph.D. students are graduate students; therefore, Chito must also be a graduate student*.

Propositional logic, also known as *sentential logic*, formally defines how such sentences are created, connected, and used for inference [58:47; 12:31; 45:166]. This earliest forms of this system are attributed to Aristotle [51:2]. His *syllogisms* contain only three sentences: a *major premise*, a *minor premise*, and a *conclusion* [58:34].

Table 10.1 lists the four forms each of these sentences can assume, and Table 10.2 illustrates several possible combinations [51:3; 58:92].

	Rule	Form
UA	Universal affirmative	Every A is B
PA	Particular affirmative	Some A is B
UN	Universal negative	No A is B
PN	Particular negative	Some A is not B

Table 10.1: *Syllogism rules*

	Example
UA	*All dogs are animals*
UA	*All collies are dogs*
UA	*Therefore, all collies are animals*
UA	*All dogs have four legs*
PA	*Some Fido is a dog*
PA	*Therefore, some Fido has four legs*
UN	*Nothing that swims is a giraffe*
UA	*All fish swim*
UN	*Therefore, no giraffe is a fish*
UN	*No dog is a human*
PA	*Some Fido is a dog*
PN	*Therefore, some Fido is not a human*

Table 10.2: *Syllogism examples*

Syllogisms have been well studied over the ages, and formal rules have been developed to prescribe their correct usage [58:36]:

- The middle term must be distributed in at lease one premise.
- A term distributed in the conclusion must be distributed in a premise.
- At most one negative premise is allowed.
- An affirmative conclusion must derive from at least one affirmative premise.
- A particular conclusion must derive from a particular premise.
- Only one particular premise may appear.
- A negative conclusion must derive from a negative premise.

Propositional logic is more flexible than syllogisms since it imposes no constraints on the number of sentences, and it supports their logical connection with *and*, *or*, *not*, *if* (⇒), and *if-and-only-if* (⇔) [45:166]. Table 10.3 lists the main inference rules [45:172; 23:493].

Rule	Application	Formula
Implication elimination	Infer a conclusion from an implication and a premise	$\frac{\alpha \Rightarrow \beta, \alpha}{\beta}$
AND elimination	Infer any of the conjuncts from a conjunction	$\frac{\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n}{\alpha_i}$
AND introduction	Infer a conjunction from a list of sentences	$\frac{\alpha_1, \alpha_2, \ldots, \alpha_n}{\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n}$
OR introduction	Infer a disjunction with anything else	$\frac{\alpha_i}{\alpha_1 \vee \alpha_2 \vee \cdots \vee \alpha_n}$
Double-negation elimination	Infer a positive sentence from a doubly negated one	$\frac{\neg\neg\alpha}{\alpha}$
Unit resolution	Infer one conjunct true if other is false in disjunction	$\frac{\alpha \vee \beta, \neg\beta}{\alpha}$
Resolution	Infer a disjunction from one false conjunct	$\frac{\alpha \vee \beta, \neg\beta \vee \gamma}{\alpha \vee \gamma}$

Table 10.3: *Common inference rules for propositional logic*

Despite its advantages over syllogisms, propositional logic is rarely used in artificial intelligence for at least the following reasons [51:357,215; 12:33; 45:185]:

- Inferences can be made only on complete sentences, not on their discrete components.
- The world model is committed only to timeless facts, so changing situations cannot be represented easily; e.g., *John was in the kitchen. John is now in the living room.*

Predicate Logic

Predicate logic, and its most common form, *first-order logic*, mitigate many weaknesses of propositional logic. The resulting logic system has become the most popular over a wide range of disciplines [4:3]. It also serves as the foundation for more advanced logics [51:41,27].

First-order logic quantifies over a closed world of objects with properties and relations [44:947; 11:95; 12:35; 51:469; 58:89]. It is a declarative approach in the sense that all operations revolve around what the objects in a model are and what they can do.

First-order logic is not without its shortcomings. One immediate observation is its non-intuitive, convoluted "mathy" representation: often the needs of well-formed logical expressions obscure the meanings and relationships of whatever is being represented. Another limitation arises in extensibility: adding new knowledge further tangles the implicit spaghetti network of interconnected expressions.

Scalability quickly becomes an issue, as well, since increasing the meaning or scope of many expressions increases their size at a much greater rate. Table 10.4 illustrates this case for simple numerical quantifiers [58:104].

Sentence	Logical Form
There are no dogs	$\neg\exists x\ dog(x)$
There is at least one dog	$\exists x\ dog(x)$
There is at most one dog	$\forall x\ \forall y\ (dog(x) \wedge dog(y)) \Rightarrow (x = y)$
There is exactly one dog	$\exists x\ dog(x) \wedge \forall y\ dog(y) \Rightarrow (y = x)$
There is at least one large dog	$\exists x\ dog(x) \wedge large(x)$
There is at most one large dog	$\forall x\ \forall y\ (dog(x) \wedge large(x) \wedge dog(y) \wedge large(y)) \Rightarrow (x = y)$
There is exactly one large dog	$\exists x\ (dog(x) \wedge large(x)) \wedge \forall y\ (dog(y) \wedge large(y)) \Rightarrow (y = x)$
There are at least two dogs	$\exists x\ \exists y\ dog(x) \wedge dog(y) \wedge (x \neq y)$
There are at most two dogs	$\forall x\ \forall y\ \forall z\ (dog(x) \wedge dog(y) \wedge dog(z)) \Rightarrow ((x = y) \vee (x = z) \vee (y = z))$
There are exactly two dogs	$\exists x\ \exists y\ (dog(x) \wedge dog(y) \wedge (x \neq y)) \wedge \forall z\ dog(z) \Rightarrow ((z = x) \vee (z = y))$
There are at least two large dogs	$\exists x\ \exists y\ dog(x) \wedge dog(y) \wedge (x \neq y) \wedge large(x) \wedge large(y)$
There are at most two large dogs	$\forall x\ \forall y\ \forall z\ (dog(x) \wedge large(x) \wedge dog(y) \wedge large(y) \wedge dog(z) \wedge large(z)) \wedge ((x = y) \vee (x = z) \vee (y = z))$
There are exactly two large dogs	$\exists x\ \exists y\ dog(x) \wedge large(x) \wedge dog(y) \wedge large(y) \wedge (x \neq y) \wedge \forall z\ (dog(z) \wedge (large(z)) \Rightarrow ((z = x) \vee (z = y)))$
There are exactly four large dogs	$\exists w\ \exists x\ \exists y\ \exists z\ dog(w) \wedge large(w) \wedge dog(x) \wedge large(x) \wedge dog(y) \wedge large(y) \wedge dog(z) \wedge large(z) \wedge (w \neq x) \wedge (x \neq y) \wedge (y \neq z) \wedge \forall v\ (dog(v) \wedge (large(v)) \Rightarrow ((v = w) \vee (v = x) \vee (v = y) \vee v = z)))$

Table 10.4: *Example quantified English sentences in first-order logic*

Furthermore, since first-order logic is so precise in its representation, it does not reflect certain ambiguity in natural language well. For example, a skit on *Saturday Night Live* once depicted a news broadcast with the following statement [4:167]:

1 a. *Every minute, a man is mugged in New York City.*
b. *We are going to interview him tonight.*

Two interpretations are possible for sentence (1a), and sentence (1b) establishes the humorous context. First-order logic would have to commit to a particular unambiguous context for sentence (1a) independent of whatever sentence (1b) contributes:

2 a. $\forall x\ minute(x) \Rightarrow \exists y\ (man(y) \wedge mugged\text{-}during(y,x))$
b. $\exists y\ (man(y) \wedge \forall x\ minute(x)) \Rightarrow mugged\text{-}during(y,x)$

Expression (2a) states that every minute there is a different man who is mugged during that time. The humorous interpretation in expression (2b) states that there is one man, and every minute, he alone is mugged! Expression (2a) does not specifically require a unique man for each minute, but in a world containing more than one man and unlimited minutes, presumably the same one would not be selected each time. Figure 10.1 illustrates two algorithms—which differ only in the position of one statement—that specify the same interpretations in procedural terms:

```
MUG_2a:                 MUG_2b:

repeat                  y ← GetMan()
  y ← GetMan()          repeat
  Mug(y)                  Mug(y)
  WaitOneMinute()         WaitOneMinute()
```

Figure 10.1: *Mugging algorithms*

Granted, such translations are difficult for any logical system, not just for first-order logic, but certain other systems provide a less rigid framework [51:414].

Higher-Order Logic

An extension to first-order logic is higher-order logic, which provides additional power and flexibility:

- It supports inference over propositional attitudes like *believe* and *know*; e.g., *Bill believes that John loves Mary* can be represented as *believes*(*Bill*,*loves*(*John*,*Mary*)). The main distinction in the order of logics comes from philosophy [34:142; 56:389; 12:74; 6:306]:
  - *de dicto* sentences make an assertion, which can be handled in first or higher-order logic.
  - *de re* sentences make an assertion about an assertion, which can handled only in higher-order logic by quantifying over predicates.
- It supports modalities of uncertainty, which are very common in natural language; e.g., *may*, *can*, *could*, *might* [51:26; 45:260]. Modal logic distills these senses into two powerful operators, which may appear in various places and in conjunction with each other:
  - The $\Diamond$ operator means *it is possible*.
  - The $\Box$ operator means *it is necessary*.

  Within temporal logic, these operators correspond to *sometimes* and *always*, respectively.

Higher-order logic exhibits other advantageous properties [51:42,27]:

- *Fewer axioms*

  Modal logic provides syntactic sugar and other niceties to first-order logic, thereby reducing the need for numerous convoluted expressions. Thus, it is considered a stronger system of inference (see page 76).
- *More natural translations*

  In using fewer axioms with more expressive power, modal logic generally makes itself more readable since each bears more resemblance to whatever it is supposed to represent. For example, the statement *it may rain* translates into modal logic with the simple $\Diamond$ operator, whereas first-order logic requires the equivalent preamble “[o]f all the states of affairs in the set of causal successors of the present, there exists at least one in which it rains” [51:42].

- *More efficient computation*

  In not having to compute over the preamble above, the implementation of a higher-order logic system may process inferences more efficiently.

Default Logic

The world exhibits considerable regularity, but it also has plenty of exceptions to the rules; e.g., birds fly, except penguins. Moreover, people have expectations based on this regularity; e.g., seeing the front of a building leads to the normally true conclusion that the building has an interior and back, etc. An exception such as a movie-studio facade is not normally considered unless evidence or suspicion brings it into play [45:459].

Knowledge representation can capitalize on this regularity with general-purpose *default* or *defeasible rules* that allow exceptions. This mechanism ties in closely with the *rules of general but not universal validity* discussed in the introduction to *Kinds of Reasoning* (see page 91) [12:97]. Default logic is non-monotonic and stable because it allows previous knowledge to be retracted or superceded without causing unmanageable contradictions [2:394; 51:373; 45:459].

Figure 10.2 illustrates the generic form of default rules [12:115; 51:376; 56:781,787; 5:119; 45:459]. If the *prerequisite* is true and the *justification* is consistent with it, then assume the *consequent*. Figure 10.3 populates this formula with the ever-popular bird example. Figure 10.4 extends it further to account for penguins.

$$\frac{\mathit{prerequisite} : \mathit{justification}}{\mathit{consequent}}$$

Figure 10.2: *Default formula*

$$\frac{\mathit{bird}(x) : \mathit{flies}(x)}{\mathit{flies}(x)}$$

Figure 10.3: *Normal bird*

$$\frac{\mathit{penguin}(x) : \neg \mathit{flies}(x)}{\neg \mathit{flies}(x)}$$

Figure 10.4: *Abnormal bird*

Defaults and generics[53] are common in linguistics, and default rules are especially useful in lexical disambiguation [56:781,789,1127; 57:174]. Abduction to achieve *proof by failure* or *negation as failure* is a common implementation [5:143; 44:1022; 2:414; 51:164].[54]

Fuzzy Logic

The section on semantics in natural language (see page 38) discussed how poorly language describes much of the world. For example, the adjective *huge* implies far greater size than *tiny* does, but a *huge mouse* is still much smaller than a *tiny giraffe*. Sowa [51:348] aptly labels this messy combination of vagueness, uncertainty, randomness, ignorance, etc. as "knowledge soup." Such issues present a huge problem in knowledge representation, and so far, only *fuzzy logic*, also known as *many-valued* or *multi-valued logic*, comes close to dealing with it adequately [12:25; 51:364; 55:666; 45:463; 44:740; 5:11].

---

[53] Objects referenced as a group, not as individuals; e.g., <u>*bear*</u> *is/are uncommon in the desert* [56:1127].

[54] Proof by failure does not necessarily prove anything related to goal unless the propositions are written well; e.g., adding an irrelevant extra clause to an expression means it may be proved [26:191; 2:414].

Unlike most logics, which have only the discrete states *false* and *true*, fuzzy logic supports operations on a continuous interval. The range, often inclusively between 0 and 1, corresponds to graded values of truth or certainty, etc.; e.g., 0 is 100% false, 1 is 100% true, any anything else lies somewhere in between [51:69]. Likewise, intervals can be applied to other vague properties, as illustrated for human age in Figure 10.5 [51:366].

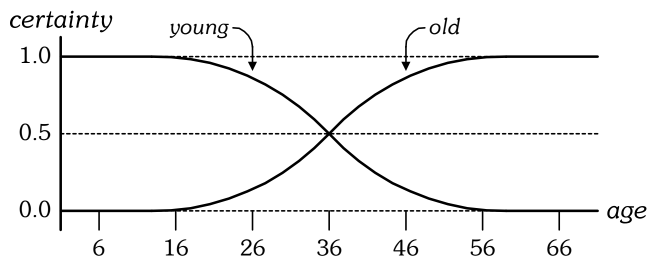


Figure 10.5: *Fuzzy function for human age*

Fuzzy logic is by no means a complete solution to the problem of vagueness. For example, the units on the x-axis in this figure are not universally applicable to everything with an age, even if the curves plausibly are; i.e., different units are needed for fruit flies, humans, sea turtles, fossils, galaxies, and so on. As usual, context plays a pivotal and confounding role.

Situation Calculus

A static world is a dead world. In the real world, most things change states, move around, and otherwise modify themselves and their surroundings. Thus, a representation of the world must be able to account for such varying situations. *Situation calculus*—and its more flexible variant, *event calculus*—provides a limited means for describing many of these aspects [48:30,33; 45:235].

The foundation of situation calculus is Leibnitz's *Principle of sufficient reason*, which claims that nothing happens without a reason [51:247,143]. States in world do not change spontaneously without grounds; rather, they conform to the expectations and limitations of the world and the valid operations that can alter them. This model is the basis of planning, which consists of three main components [45:341]:

- *Initial state*

  Every plan has a starting point; e.g., a person crosses a room first by entering the room.

- *Goal state*

  Every plan should have an ending point;[55] e.g., the room has been crossed when the person reaches the other side.

- *Operators*

  Every state between the initial state and goal state can be transitioned according to axioms and constraints in the model; e.g., crossing a room is most likely done on foot and may occur in a

[55] The goal need not be fully known in advance or ever reached. Research and wars on terrorism are protracted examples that are considered over at some arbitrary point, regardless of whether any "goal" was achieved.

beeline fashion or randomly, etc. Flying across a room is not within the capabilities of a person in this context.

One terribly vexing problem in this representation is how to update the states in a model after each transition [44:949; 51:245; 26:166]. This so-called *frame problem* is thought to be the hardest cognitive challenge in artificial intelligence [34:81,267]. Picture the following scenario:

> In a military arsenal are hundreds of weapons, thousands of rounds of ammunition, tons of high explosives, several drums of rocket fuel, as well as a ball of yarn and a lit candle on a ledge over a pan of gunpowder. There is also a little white kitty.

If the kitty knocks the ball of yarn into the pan, what happens to the states of everything in the arsenal? Realistically, only the yarn will be affected; i.e., its position will be different, and it may have unraveled somewhat. However, if the kitty knocks the candle into the pan, the states will require some massive updating, to say the least! The moral of this story is that slight differences in a single action may have entirely different explicit and implicit consequences[56], and any representation of such a model must account for them.

The frame problem manifests itself in several sinister forms [45:207; 12:200]:

- *Inferential frame problem*

  The transition between states changes some things and leaves others the same; e.g., taking a step forward toward the middle of a room means a person is still in the room, whereas what he or she can see from the new vantage point may differ.

- *Representational frame problem*

  Axioms must account for every valid transition between states; e.g., moving forward in a room normal entails taking a step forward, but a kung fu leap may have the same result.

- *Qualification problem*

  Axioms must account for correct actions that fail for an infinite number of reasons; e.g., crossing a room may not be prudent if it is on fire or a psychopath with a chainsaw is present.

- *Ramification problem*

  Every action has implicit consequences that must be inferred between states; e.g., a person keeps the same lint is his or her pocket while crossing a room, but the oxygen molecules in it will change.

[56] Mathematical theories of *chaos* hypothesize a connection between the flapping of butterfly wings in South America on global weather patterns!

Section 11

Mechanisms of Reasoning

Procedural Approaches

The merits of declarative and procedural knowledge representations received intense debate by the artificial-intelligence community in decades past. The unsurprising result is now a consensus that each representation has its place, and in fact, combinations of both, known as *hybrid approaches*, are actually necessary [44:1017; 37]. The previous section discussed declarative approaches as the form of a solution rather than as the mechanisms to attain it. Procedural approaches consider the opposite—what the steps to the goal are [40:435].

For comparison, consider a naive sort implemented in declarative and procedural (imperative) programming paradigms:

```
DECLARATIVE_SORT:
list_sorted is any permutation P of list_unsorted where each element Ei
in P is greater than Ei+1

PROCEDURAL_SORT:
repeat until list_unsorted is empty
   find the largest element E in list_unsorted
   remove E from list_unsorted
   append E to list_sorted
```

Both algorithms return a sorted list. The declarative approach describes the properties of a sorted list and effectively says to the computer: do whatever is necessary until one is found. The procedural approach instead specifies the exact steps toward the same goal. In programming, the advantages of each paradigm are clear: declarative programs tend to be smaller, more readable, more extensible, and more verifiable, while procedural programs tend to be more efficient [12:3; 46:541]. Indeed, the declarative sort runs in $O(n!)$ time versus $O(n^2)$ for the procedural—an indescribably gargantuan difference![57]

[57] For example, to sort 50 elements takes at most roughly 3^{64} iterations versus 2,500 iterations, respectively!

Luckily, the differences in efficiency between the two approaches for knowledge representation are much less extreme. Consequently, the choice of one over another often depends more on the needs of the application.

Procedural approaches operate under a common framework known as a *production system* [34:104]. Within this framework, *production rules* define knowledge in terms of steps toward some goal. The generic form is

production-rule = condition + action

where *condition* specifies what must hold true for *action* to be executed [34:104,135].

Scripts

The world is a chaotic and unpredictable place, but within this apparent disorder can often be found considerable regularity. For instance, going out to eat at a restaurant entails countless details that can vary widely or may not be present at all. The basic steps and their typical order, however, are relatively fixed:

1. Go to the restaurant.
2. Order the food.
3. Receive the food.
4. Eat the food.
5. Pay for the food.
6. Leave the restaurant.

Granted, many exceptions or elaborations to this framework are possible; e.g., at a fast-food restaurant, Step 5 usually occurs immediately after Step 2 or Step 3. Nevertheless, eating out can arguably be distilled into just six high-level steps. Furthermore, each step can be decomposed into substeps; e.g., Step 5 involves a transaction where money is given (and in most cases, a tip) and possibly change is received.

Such a stepwise description of events to achieve a goal is known as a *script*. Although a generic script rarely overlays an actual sequence of events exactly, many of its steps are usually present. This structural similarity is especially useful for computers to establish context and to understand the motivation and actions of the agents within it [34:195]. Narratives are notoriously sparse with explicit details, and, as it has been shown throughout this paper, computers are notoriously bad at inferring implicit details [34:193]. Scripts fill in the gaps somewhat.

Although scripts are clearly useful, two difficulties arise in manipulating them programmatically [2:477; 34:197]:

- *How does the system know which scripts are relevant?*

  Any real-world NLP system needs scripts for the many conceivable scenarios it may encounter. Realistic scripts have far more steps in greater detail than is shown above, and the exceptions, elaborations, and so on must be accounted for.

For a script to be useful, it must also be selected as appropriate for the given scenario. The mechanism behind this selection is complex and often unreliable. Furthermore, selecting the incorrect script could be more damaging than having no script at all.

- *How does the system keep track of where it is in the script and its embedded subscripts?*

  Scripts are compositional and thus form a tangled stack-like structure as they play out. Maintaining the current step in light of the many tangential branches is difficult. Furthermore, selecting the next step incorrectly could lead a system astray with little chance for recovery.

These deficiencies, among others, have led to the development of additional narrative representations that build upon each other hierarchically [51:293; 34:198]:

- *Scenes*

  If a high-level script like EAT-AT-RESTAURANT above cannot be matched with the current scenario, then related lower-level processing is obviously not applicable. Scenes encode abstract commonalities of basic models and do not deal with exceptions or elaborations. They can be used as triggers to determine which scripts are relevant.

- *Memory organization packets (MOPs)*

  A collection of scenes can be interconnected with respect to their triggers. If an enabling scene cannot be found, then scenes dependent on it can be safely ignored. The notion of GOING-OUT-TO-EAT would be considered a MOP. Within it might be found the EAT-AT-RESTAURANT scene, as well as a GOING-TO-A-DRIVE-THROUGH scene, a GOING-FOR-TAKE-OUT scene, and so on. The overall scenario revolves around obtaining prepared food; the individual scenarios focus on how this goal is achieved.

- *Thematic organization packets (TOPs)*

  A collection of MOPs describes high-level, very abstract notions. For instance, the goal of OBTAINING-FOOD can be satisfied by GOING-OUT-TO-EAT, by HAVING-FOOD-DELIVERED, by COOKING-AT-HOME, and so on.

The formal languages that define scripts are based on approaches discussed throughout this paper, so they are not addressed here. One notable exception is Petri nets, which are especially useful for event-driven simulations [51:235]. This graph-based implementation relies on various mechanisms used in computer operating systems and networks to manage the lock-and-key structure of interdependent triggers; e.g., synchronous and asynchronous data transfer, message passing, semaphores, etc. [44:1402; 51:141].

Expert Systems

Experts in almost any subject are valued because they know their material far better than the average person does. Artificial-intelligence applications that attempt to emulate the abilities of experts are known as *expert systems*. Such systems automate and improve the familiar troubleshooting guides available for many limited problem domains like auto repair and software support. They also extend into highly specialized domains like medical diagnosis.

Table 11.1 shows a subset of an automotive troubleshooting guide related to instrumentation problems [17:504]. Such a simple representation provides only an overview of a small range of conceivable problems. As such, it is intended mainly for people of average ability and resources; professional mechanics have a far more detailed version, which they use in conjunction with their expertise and advanced tools. Nevertheless, this representation could be considered a bare-bones expert system since it embodies the knowledge and experience of whoever wrote it.

Symptom	Cause	Remedy
All gauges do not operate	Blown fuse	Replace gauge main fuse
	Defective instrument regulator	Replace regulator
All gauges read low or erratically	Dirty instrument voltage regulator	Clean regulator
	Defective instrument voltage regulator	Replace regulator
All gauges are pegged	Loss of ground between instrument voltage regulator and frame	Fix break in connection
	Defective instrument regulator	Replace regulator

Table 11.1: *Subset of automotive troubleshooting guide*

True expert systems, which are interactive and provide more detailed information, are based on three main components [44:686,1438; 51:156]:

- *Language*

  Straightforward `if-then` rules are typically employed, often taking into account probability or uncertainty. User interactivity allows conditional expressions to narrow the scope of problem solving. Decision trees, which will be discussed in the next section, can also play a role.

- *Knowledge base*

  This repository serves as the embodiment of relevant contributions from experts. It usually contains two kinds of knowledge:

  - *factual knowledge* describes common details, assumptions, prerequisites, etc., which are agreed on by most experts.
  - *heuristic knowledge* incorporates less tangible aspects of problem solving such as rules of thumb, ad hoc approaches, judgment calls, and other experience that may differ or even conflict among experts.

- *Inference or reasoning engine*

  Interactive execution of the language on the knowledge base is supposed to generate relevant and helpful output. The mechanisms behind this process differ by implementation, but they are based on the approaches discussed throughout this paper. In addition to evaluating conditional rules, forward and backward chaining (see page 94) are commonly used.

Expert systems, like everything in the field of knowledge representation, vary widely according to different applications, domains, implementations, etc. In spite of such variation, all systems are based on the general architecture shown in Figure 11.1 [45:687].

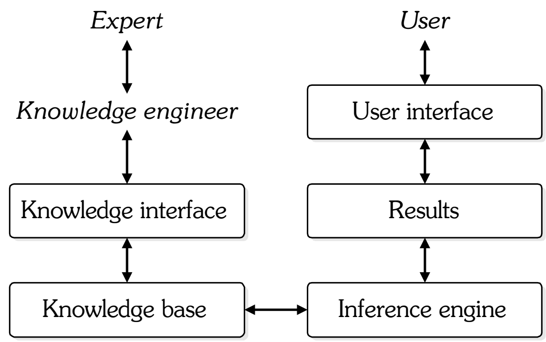


Figure 11.1: *Basic expert-system architecture*

Section 12

Mechanisms of Reasoning

Graph-Based Approaches

Graphical approaches to knowledge representation lie somewhere between declarative and procedural approaches and borrow many features from both. In fact, in many cases, they are simply the semantically equivalent visual counterpart to text forms and, as such, can be freely transformed back and forth [45:317].

Graphical models have at least two practical advantages over linear ones [49; 30; 45:317]:

- *Readability*

  A picture is worth a thousand words. The same holds true for graphical representations, which have been shown for visual programming languages to exhibit several useful properties:

  - evidence suggests that abstract reasoning is pictorial in nature.
  - multidimensional representations may support visual reasoning better than linear forms.
  - visual representations lend themselves well to the psychological notion of chunking.
  - images and thought may be transformations of each other.
  - experiments indicate that the mind stores sensory information in very detailed form.
  - graphical models usually relate in form to the task they represent, whereas linear models may bear no resemblance.

- *Implementation*

  Graph theory is well established in mathematics and computer science. Some algorithms are inherently more efficient on graphs.

Inheritance

Throughout time, philosophers, biologists, astronomers, and many others, have observed an orderly structure to almost everything. The universe, for example, contains galaxies, which in turn contain solar systems that have a sun and perhaps planets, and so on. From largest to smallest, most general to most specific, etc., the world lends itself well to taxonomic description.

Knowledge representation can benefit from this structure through hierarchical graphs. Each node is some element in the model being represented. It contains both properties that describe it and various kinds of links to other related nodes. Inheritance is performed by following the links and recording the properties and relations encountered along the way [34:92]. For example, Figure 12.1 describes a *dog* as a *canine*, which is a *mammal*, and hence a *living-thing*. If a *living-thing* has the property of *being mortal*, then by inheritance, so do *mammal*, *canine*, and *dog*. Likewise, if *mammal* has the property of *breathing air*, then so do *canine* and *dog*, etc.

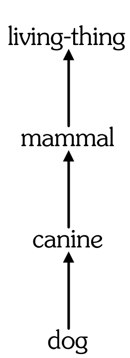


Figure 12.1: *dog*

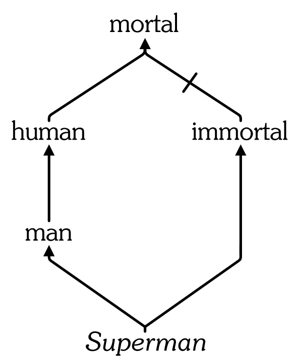


Figure 12.2: *Superman*

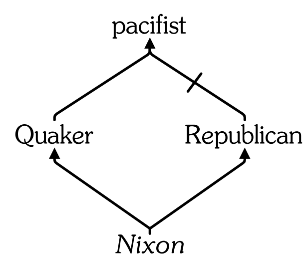


Figure 12.3: *Nixon*

An inheritance hierarchy is constructed in one of two ways [45:320; 51:151,381]:

- *Single inheritance*

  A node can have only one parent per relation; e.g., *dog* **is-a** *canine*.

- *Multiple inheritance*

  A node can have multiple parents per relation; e.g., *dog* **is-a** *canine* and **is-a** *house-pet*. Since the properties of each parent are inherited, a potential for conflict exists, as shown in Figures 12.2 and 12.3.

Furthermore, inheritance can abide by different rules [56:793,812; 5:248]:

- *Strict taxonomic inheritance*

  Only monotonic inheritance is supported, and all nodes must be taxonomically related; e.g., *dog* **is-a** *mammal* **is-a** *animal* and so on. Property relations like **has-color** are not allowed.[58]

- *Mixed or defeasible taxonomic inheritance*

  Non-monotonic inheritance is supported, but all nodes must still be related taxonomically.

[58] Property relations through convoluted means may be possible; e.g., *Superman* **is-a** *immortal-thing*. However, such trickery likely undermines the validity and usefulness of the taxonomy. Known as *ontological promiscuity*, it opens the door for unrestricted reification of properties into objects [57:57; 45:258].

- *Skeptical inheritance*

  Inheritance may not draw unique conclusions due to conflicts, as shown in Figures 12.2 and 12.3. In such case, other reasoning may be applied to make the appropriate choice; e.g., if Quakers have stronger beliefs than Republicans do, then their branch should exhibit greater influence.

  Two mechanisms that help cope with conflicts are based on AND/OR trees discussed below [5:249]:

  - *credulous reasoning* requires at least one branch to lead to a non-conflicting conclusion.
  - *ideally skeptical reasoning* requires all branches to lead to a non-conflicting conclusion.

Taxonomic inheritance can be extended to support relations beyond **is-a** at the expense of additional complexity and potential for conflict [56:796,800].

Search

Almost all artificial-intelligence applications search a data structure for a predefined solution [2:600]. This structure, often some kind of tree, contains intermediate states and desired goals. How a search proceeds from state to state toward a goal is based on decisions made at each state. The mechanisms behind these decisions range from straightforward conditions like $\texttt{cost(state}_1\texttt{)} < \texttt{cost(state}_2\texttt{)}$ to loosely defined heuristic rules like *odd states tend to work better* [51:245; 44:90; 45:94].

Trees

Trees are mathematical models and computer data structures that have been studied extensively in many fields. No background information is provided here, as it can be found in any computer-science textbook on data structures.

Decision Trees

Propositional knowledge (see page 105) lends itself well to *machine learning* [45:31; 11:642]. This subject is outside the scope of this paper, but one common approach, *decision trees*, warrants brief discussion because it overlaps well with search.

A common representation of datasets for human consumption is a table. Most people are familiar with looking up values in columns and reading across the appropriate row to find desired information. Table 12.1 presents a simplistic example for deciding whether to go hiking based on the weather. This table can be translated directly into a Boolean function as a disjunction of row functions; e.g.,

```
hike = (sky( clear ) ∧ rain( none )) ∨
       (sky(broken) ∧ rain( none )) ∨
       (sky(broken) ∧ rain(light)) ∨
       ...
```

Moreover, the table can also be translated into a tree, as illustrated in Figure 12.4.

Sky	Rain	Hike?	Recommendations
Clear	None	Yes	None
Broken	None	Yes	None
Broken	Light	Yes	Bring umbrella or raincoat
Broken	Heavy	Yes	Bring umbrella and raincoat
Overcast	None	Yes	Leave camera at home
Overcast	Light	No	None
Overcast	Heavy	No	None

Table 12.1: *Decision table*

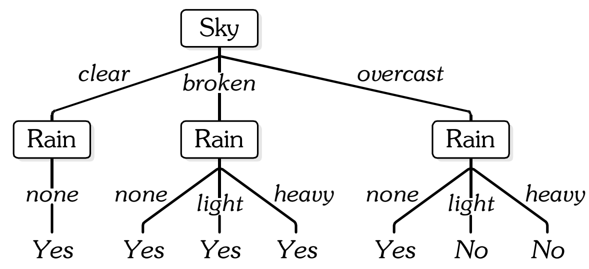


Figure 12.4: *Decision tree*

A decision tree is traversed downward from the root. The question at each node is evaluated based on the available information, and the appropriate branch is followed recursively. When a leaf is reached, its value is returned as the goal.

Expert systems (see page 115) benefit from decision trees for several reasons [45:538; 55:526]:

- *Interactivity with the user*

  Each node presents a question that is directly relevant to its branch, and traversing the branch narrows subsequent questions.

- *Justification at each step*

  As each question is answered, information can be generated to explain the solution when it is finally reached, as indicated in the recommendations column of Table 12.1. An extended example is: *do not go hiking because an overcast sky usually gets worse, and heavy rain often results in flooding. Plus photographs do not come out nice.*

- *Flexibility with missing information*

  If an exact question or answer is not available, nearby branches, which are presumably somehow related, could be investigated. An exact goal may not be reached, but helpful information could be produced. In the event that no related branch can be found, the justifications leading up to the point of failure may still provide some clues in problem solving.

AND/OR Trees

Traversing a decision tree involves following only one branch per visited node. There is no mechanism to handle multiple selections except for, say, running an expert system again, or bookmarking a particular question for revisiting. A more advanced tree representation allows multiple questions to be connected with logical functions [55:146,678,694]:

- *AND node*

  All matching branches under a node are traversed, and *all* must return a goal.

- *OR node*

  All matching branches under a node are traversed, and *at least one* must return a goal.

Combining the goals into a useful solution is tricky. Typically AND/OR trees are used for tasks with well-defined nodes, branches, and goals; e.g., games and puzzles [55:678]. Furthermore, specialized search techniques beyond the scope of this paper are needed [*ibid*].

Searching Trees

A decision tree combines both a representation and a search strategy. Normally these two aspects are addressed separately to consider the merits of different approaches; i.e., which type of tree is most appropriate for the application, and which search strategy works best on it? The details of other tree representations are beyond the scope of this paper, so only search strategies will be discussed. The main properties of any strategy relate to its computability [45:73]:

- *Completeness*
  The search is guaranteed to return a solution if one exists.
- *Optimality*
  The search is guaranteed to return the best solution if more than one exists.
- *Time complexity*
  The search will take a known maximum amount of time based on the size of the tree.
- *Space complexity*
  The search will take a known maximum amount of memory based on the size of the tree.

Other properties involve how the search proceeds through the tree toward a goal [45:73; 55:676]:

- *Uninformed or blind search*
  No information is known a priori about the number of steps or the cost of searching for a goal. The search only knows whether it has found a goal or whether it must continue searching.
- *Informed or heuristic search*
  Information is known a priori about the nature of the search space, and it can be used at each step to direct the search in presumably the best direction toward a goal.

The following strategies form the backbone of search [45:85; 55:694,697; 2:603]:

- *Breadth-first search*
  Each node on a level of the tree is visited horizontally before proceeding to the next lower level. This strategy is complete, as well as optimal if all paths have the same cost. However, the large time and space complexities limit its practical use.
- *Depth-first search*
  Each branch is visited vertically until a goal or leaf is reached. A leaf results in traversal returning to the original branch and visiting the next branch. This strategy is neither complete

nor optimal. Furthermore, the time complexity in large or infinitely deep trees limits its practical use.

An alternative form, *depth-limited search*, imposes a fixed cutoff point in the tree below which no further searching can be done. It is both complete and optimal only if the goal lies within the available search space.

Iterative deepening uses a dynamically adjustable cutoff point in the same way. It is both complete and optimal.

- *Best-first search*

  The branches to traverse at each node are determined by a heuristic rule that presumably leads the search in the best direction. This strategy may use either breadth-first or depth-first searching.

- *Other searches*

  Similar strategies have been developed to exploit the advantages and reduce the weaknesses of these basic strategies above; e.g., *uniform-cost*, *bi-directional*, *A**, *branch and bound*.

Semantic Networks

Research suggests that human knowledge is organized hierarchically in a mental network of related knowledge [51:4; 11:97]. Semantic networks abstract this notion into a graph-based implementation that has long been used in philosophy, psychology, and linguistics, as well as more recently in artificial intelligence and natural language processing [50].

A semantic network generally consists of concept nodes with relational connections. Following the connections as they fan out from a concept reveals its compositional structure and semantic properties. Figure 12.5 illustrates a rudimentary semantic network describing a robin as a bird that eats worms and has red feathers.

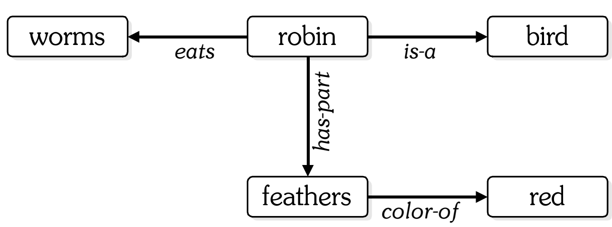


Figure 12.5: *Semantic network for 'robin'*

Six kinds of semantics networks are most common [50]:

- *Subsumption or definitional networks*

  Subtypes and **is-a** relations are emphasized between concepts and subconcepts that are assumed true within the represented domain. The hierarchical structure supports powerful and efficient inheritance. This structure, commonly used in ontologies (see page 127), closely parallels semantic memory (see pages 71 and 83).

- *Assertional networks*

  Propositions are interconnected. They are assumed contingently true unless evidence to the contrary can be found or inferred. This structure closely parallels episodic memory (see pages 71 and 84).

- *Implicational networks*

  Nodes are interconnected by logical implication to form chains of causality, belief, or inference.

- *Executable networks*

  Operations are performed as nodes are visited. Petri nets (see page 115) are an example.

- *Learning networks*

  Network interconnections are built or extended through training on examples. Artificial neural networks (see page 129) are an example.

- *Hybrid networks*

  Any of these networks can be combined.

One ironic criticism of semantic networks addresses their often-unclear semantics: relations can be defined arbitrarily, and care must be taken to avoid proliferating unnecessarily similar, redundant, and useless ones, etc. [5:191; 34:105]. Another criticism takes aim at how semantic networks tend to focus more on relations within the network than between it and the domain being represented [34:105]. Despite these shortcomings, carefully developed networks have proved successful for many applications. Highly popular *WordNet*, for instance, maintains a huge structure of rich interrelations between English words [16; 11:631].

Frames

An alternative form of semantic networks is a frame-based representation. Instead of the relatively incoherent tangle of connected nodes in semantic networks, frames organize details related to each concept into clusters [2:400]. Each relation is represented as a *slot*, and its value is a *filler*, which can take an atomic value, a function, or a reference to another frame [34:133; 54:193]. Figure 12.6 extends the example from Figure 12.5 into a network of frames.

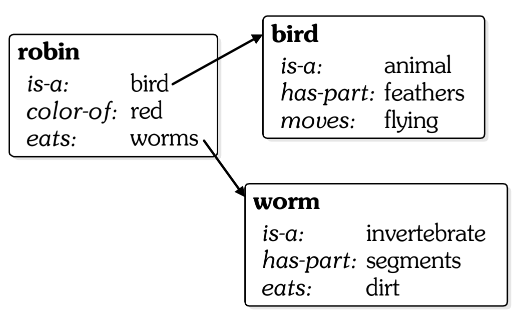


Figure 12.6: *Frames for 'robin'*

Conceptual Graphs

Aspects of semantic networks and frames can be combined into a *conceptual graph*. This representation is as logically precise and expressive as first-order logic, but it lends itself far better to human readability and computational tractability [51:23,476; 49; 47]. In addition, it provides mechanisms for many difficult linguistic constructions like quantification and indexical relations (see pages 46 and 48, respectively) [51:476].

Conceptual graphs are built from two components interconnected by directed arcs [51:23,476]:

- *Concepts*

  Objects and properties appear in concept boxes with two slots:

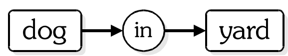
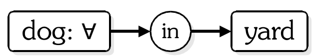
  - a *type*, which corresponds to the notion of type or object class in programming languages
  - a *referent*, which generally plays the role of variables in programming languages. It may be a combination of a named instance, a variable, a quantifier, or metalanguage features

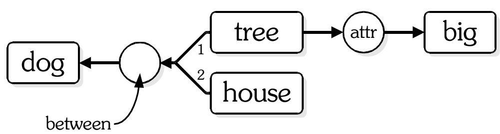
  The type is always required, whereas the referent is optional. If it is not present, the default is existential quantification; e.g., `[dog]` specifies some unnamed dog, whereas `[dog: Rambo]` specifies a particular dog.

- *Conceptual relations*

  Concepts are interconnected by relation circles, which function analogously to prepositions in natural language. Each relation maps a number of inputs to an output; e.g., *between* take two inputs, whereas *in* takes one or more; both return a spatial-relation output.

Table 12.2 illustrates several translations of natural language into two conceptual-graph representations [49]. Graphs involving advanced features like propositions, situations, coreference links, and beliefs would appear in similar form; the only additions are dashed lines and boxes around subgraphs.

Graph Form	Linear Form
A dog is in a yard 	`[dog: *x] [yard: *y] (in ?x ?y)`
Every dog is in a yard 	`[dog: @every*x] [yard: *y] (in ?x ?y)`

A dog is between a big tree and a house

	`[dog: *w] [house: *x] [tree: *y]`
	`[big: *z] (between ?x ?y ?w)`
	`(attr ?y ?z)`

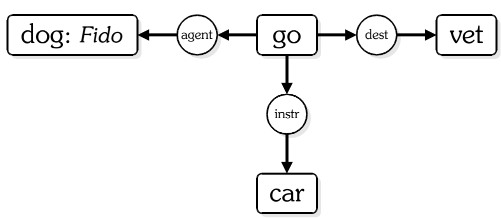
The dog Fido is going by car to a vet

	`[go: *w] [dog: 'Fido' *x] [vet: *y]`
	`[car: *z] (agent ?w ?x) (dest ?w ?y)`
	`(instr ?w ?z)`

Table 12.2: *Simple conceptual graphs*

Ontology

The union of semantic networks, frames, and conceptual graphs, along with other approaches, produces an ontology. This rich knowledge representation with vast potential carves up the world—or more realistically, a subset of a domain—as a complex subsumption hierarchy of frame-like components [54:184; 55:719]. Moreover, it is not just a representation of knowledge; it can also be viewed as… [21]

… a philosophical discipline
… an informal conceptual system
… a formal semantic account
… a specification of conceptualization
… a vocabulary used by logical theory
… a meta-level specification of logical theory
… a representation of conceptual systems through logical theory

Natural language processing benefits from these properties in so many ways that "[i]n the field of NLP, there is now a consensus that all NLP systems that seek to represent and manipulate meanings of texts need an ontology" [33:6]. In particular, an ontology… [33:5]

… provides a way to represent the meaning of text in a language-neutral way
… allows lexicons for different languages to share common knowledge
… enables source-language analyzers and target-language generators to share knowledge
… stores selectional restrictions and other pieces of world knowledge
… fills gaps in meaning by inferences based on its representation of conceptual knowledge
… helps resolve semantic and pragmatic ambiguity
… serves as a classification of people, places, social roles, organizations, etc.
… maintains a repository of selectional preferences on the composition of meaning
… supports inferences on the topology of the network to determine semantic relatedness

Since as far back as Aristotle's time, people have been trying to define a formal structure of the world in terms of what its objects are, how they are related, what they do, and so on [2:231]. Representations differ widely, and no single one could ever be considered indisputably correct in all respects.[59] What they generally have in common, however, is a semantic structure based on variations of **is-a** and **part-of** relations [26:125; 2:307; 44:949; 53:14]. Concepts in an ontology are often linked taxonomically from more general to more specific, as well as lattice-wise with other interrelations. Taxonomical subbranches of a node represent *differentiae*, which are the properties or features that distinguish them from their parent node and each other; e.g., *man* and *woman* are both more specific subtypes of *human*, but each is so in slightly different ways [51:494].

The form and contents of an ontology are a matter of great debate. However, since the goal is usually to represent the real world, it is not surprising that the range of paradigms is limited [42; 22; 39:70; 45:228; 34:191; 56:898]:

- *Ontology oriented toward objects*

  Nodes describe tangible or definable entities in world, such as physical objects, events, regions, quantities of matter, and so on.

- *Ontology oriented toward concepts*

  Nodes describe meta-level categories used to model the world, such as abstract concepts, properties, qualities, states, roles, parts, intentions, plans, beliefs, and so on.

Further distinctions can be made by the domain an ontology represents [42; 22]. A *domain-independent* model is neutral with respect to any particular application, task, language, culture, etc. A *domain-dependent* or *situated* model, on the other hand, is designed with a specific goal in mind.

[59] See the section of semantic features on page 42 for a related discussion.

Figure 12.7 depicts in abridged form the top-level structure of the *Mikrokosmos* ontology [33:12].

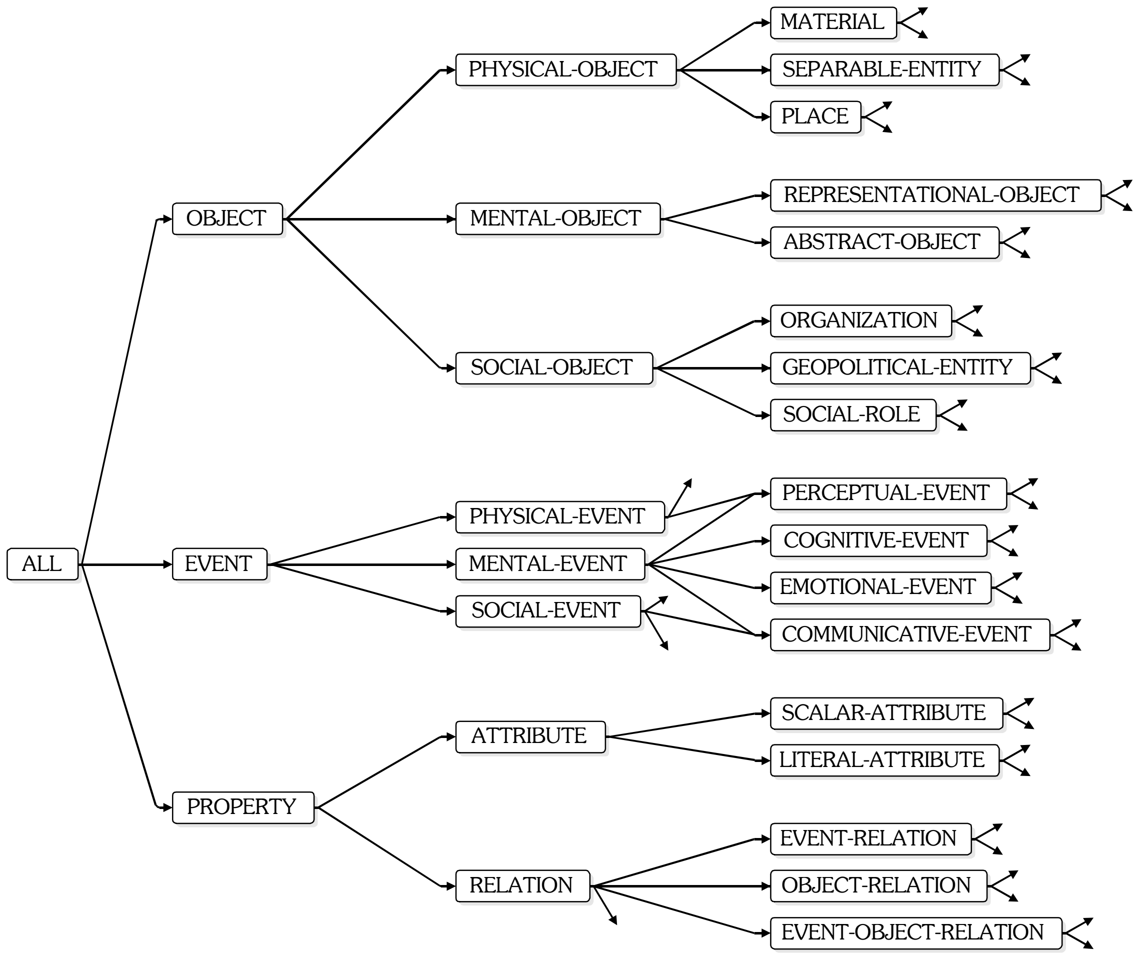


Figure 12.7: *Mikrokosmos top-level ontology*

Artificial Neural Networks

Neural networks[60], also known as *connectionist models*, are plausible representations of the way the brain works as a collection of neurons and synaptic connections [24:6]. Many successful applications in engineering and artificial intelligence employ them; e.g., pattern recognition, robotic vision, data mining, etc. Natural language processing may benefit as well.

60 "Artificial" is typically omitted when no confusion would ensue.

Neural architectures vary widely, but they all interconnect functional units in some way. Figure 12.8 illustrates a rudimentary switching unit that is duplicated as circles throughout the generic network in Figure 12.9 [63:115]. It takes any number of weighted inputs and passes their sum through a discrete or continuous function to produce an output.

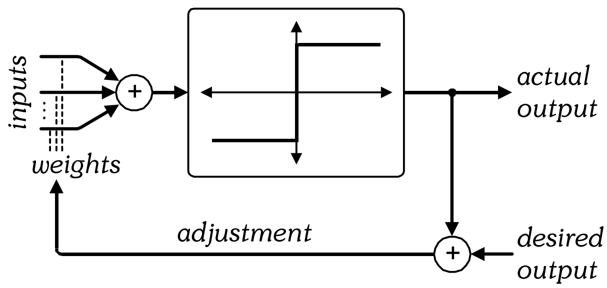


Figure 12.8: *Switching unit*

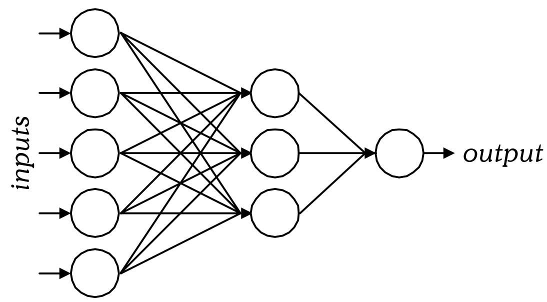


Figure 12.9: *Feedforward network*

Before a network can be used on real data, it must be trained on test data. Training is either *supervised* or *unsupervised* depending on whether or not the network has the correct answers available, respectively [24:65,57; 44:1057; 45:528]. For training, if the actual output does not match the desired output, iterative adjustments are made to the input weights automatically and training continues. Training succeeds once the network stabilizes with no further iterations. For execution, the adjustment connections are disabled and the actual output is used as the result.

Neutral networks have their place in NLP systems, but alone they cannot serve all the various linguistic functions. As a result, they are normally coupled in one of three ways with conventional symbolic approaches [11:825]:

- *Loosely coupled architecture*

  Connectionist and symbolic models are separate components and the latter provides input into the former.

- *Tightly coupled architecture*

  Connectionist and symbolic models are separate components and each provides input into the other.

- *Fully integrated architecture*

  Connectionist and symbolic models are a single component.

Neural networks offer many advantages in wide range of tasks [45:566,583,625; 11:672; 24:4]:

- *Expressiveness*

  Although connectionist models do not have the expressive power of general logical representations, they nevertheless perform very well on many kinds of input.

- *Generalization*

  Flexibility and expressiveness allow connectionist models to adapt to a wide variety of tasks.

- *Computational efficiency*

  Networks are inherently parallel, so fully trained connectionist models operate very efficiently.

- *Plausible biological model*

  The tangled interconnections between neural components are believed to be more analogous to the way the human mind functions than symbolic models are.

- *Noise tolerance*

  Strong generalization allows connectionist models to smooth over anomalous or spurious data.

- *Graceful degradation*

  As the quality of the inputs deteriorates, the quality of the outputs tends to decline gradually.

On the negative side, neural networks suffer from several limitations [11:674; 45:566,584]:

- *Scalability*

  Connectionist models that work on small datasets may not extend to larger datasets.

- *Transparency or inscrutability*

  As a black-box implementation, the inner workings of connectionist models are unreachable. Debugging a network or tweaking its performance is extremely difficult.

- *Prior knowledge*

  Selecting an appropriate connectionist model and configuring it to a particular task is a poorly defined task that relies heavily on experience, perseverance, and luck.

- *Structure representation*

  Natural language consists of many compositional symbolic structures that do not lend themselves well to subsymbolic representations or processing.

□

References

[1] Abate, F., et al., eds., *Oxford Dictionary and Thesaurus*, Oxford University Press, New York, 1996.

[2] Allen, J., *Natural Language Understanding*, Benjamin/Cummings, Redwood City, CA, 1995.

[3] Allen, J., "Maintaining Knowledge about Temporal Intervals." Communications of the ACM, November 1983, pp. 832-843.

[4] Barwise, J, Etchemendy, J., *Language of First-Order Logic*, CSLI Publications, Stanford, CA, 1992.

[5] Brewka, G., *Principles of Knowledge Representation*, CSLI Publications, Stanford, CA, 1996.

[6] Chierchia, G, McConnell-Ginet, S., *Meaning and Grammar: An Introduction to Semantics*, MIT Press, Cambridge, MA, 2000.

[7] Chomsky, N., *Syntactic Structures*. Janua linguarum, Ser. minor, Vol. 4. Mouton, Gravenhage, 1957.

[8] Cowie, J., "The treatment of collocations and idioms in learner's dictionaries" Applied Linguist 2:223-235, 1981.

[9] Crystal, D., *Cambridge Encyclopedia of Language*, Cambridge University Press, Cambridge, England, 1987.

[10] Crystal, D., *Cambridge Encyclopedia of the English Language*, Cambridge University Press, Melbourne, Australia, 1996.

[11] Dale, R., Moisl, H., Somers, H., eds., *Handbook of Natural Language Processing*, Marcel Dekker, New York, 2000.

[12] Davis, E., *Representations of Commonsense Knowledge*, Morgan Kaufmann, San Mateo, CA, 1990.

[13] Davis, R., Shrobe, H., Szolovits, P., "What is Knowledge Representation?" AI Magazine, 14(1):17-33, 1993.

[14] Dewey, G., *English Spelling: Roadblock to Reading*, Teachers College Press, New York, 1971.

[15] Doyle, P., "Natural Language AI Qual Summary" Web Page: http://www.cs.dartmouth.edu/~brd/Teaching/AI/Lectures/Summaries/natlang.html

[16] Fellbaum, C., ed., *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA, 1998.

[17] Freeman, K., ed., *Jeep Wagoneer/Commando/Cherokee/Truck 1957-86*, Chilton, Radnor, PE, 1986.

[18] Fung, P., Wu, D., "Statistical augmentation of a Chinese machine-readable dictionary" Proceedings of Second Workshop on Very Large Corpora (WVLC-94), 1994.

[19] Green, G., Morgan, J., *Practical Guide to Syntactic Analysis*, CSLI Publications, Stanford, 1996.

[20] Green, G., *Pragmatics and Natural Language Understanding*, Lawrence Erlbaum, Hillsdale, NJ, 1989.

[21] Guarino N., Giaretta P., "Ontologies and Knowledge Bases: Towards a Terminological Clarification" In N. J. I. Mars (ed.), Towards Very Large Knowledge Bases, IOS Press, 1995.

[22] Guarino, N. Poli, R., eds., "Formal Ontology, Conceptual Analysis and Knowledge Representation" Special Issue of the International Journal of Human and Computer Studies, 43:5/6.

[23] Hausser, R., *Foundations of Computational Linguistics*, Springer, Italy, 1999.

[24] Haykin, S., *Neural Networks: A Comprehensive Foundation*, Macmillan Publishing, Englewood Cliffs, NJ, 1994.

[25] Hobbs, J., "Resolving pronoun references" Lingua 44:311-338, 1978.

[26] Iwanska, L, Shapiro, S., eds., *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language*, AAAI Press, Menlo Park, CA, 2000.

[27] Jannedy, S., Poletto, R., Weldon, T., eds., *Language Files*, Ohio State Press, Columbus, 1994.

[28] Kay, P., *Words and the Grammar of Context*, CSLI Publications, Stanford, CA, 1997.

[29] King, T., *Configuring Topic and Focus in Russian*, CSLI Publications, Stanford, 1995.

[30] Kremer, R., "Visual Languages for Knowledge Representation" Eleventh Workshop on Knowledge Acquisition, Modeling and Management, April 18-23 1998.

[31] Larson, R., Segal, G., *Knowledge of Meaning: An Introduction to Semantic Theory*, MIT Press, Cambridge, MA, 1996.

[32] Levin, B., *English Verb Classes and Alternations*, University of Chicago Press, Chicago, 1993.

[33] Mahesh, K., "Ontology Development for Machine Translation: Ideology and Methodology" Technical Report MCCS-96-292, Computing Research Laboratory, New Mexico State University, 1996.

[34] Markman, A., *Knowledge Representation*, Erlbaum, New Jersey, 1999.

[35] McArthur, T., ed. *Oxford Companion to the English Language*, Oxford University Press, New York, 1992.

[36] McCutcheon, M., *Roger's Super Thesaurus*, Writer's Digest, Cincinnati, OH, 1995.

[37] Minsky, M., "Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy" AI Magazine 12(2):34-51, 1991.

[38] Neufeldt, V., Guralnik, D., et al., eds., *Webster's New World Dictionary of American English*, Simon & Schuster, New York, 1988.

[39] Nirenburg, S., Carbonell, J., Tomita, M., Goodman, K., *Machine Translation: A Knowledge-Based Approach*, Morgan Kaufman, San Mateo, CA, 1992.

[40] Norvig, P., *Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp*, Morgan Kaufmann, San Mateo, CA, 1992.

[41] Pinker, S., *The Language Instinct: How the Mind Creates Language*, Harper Collins, NY, 1994.

[42] Poli, R., "Framing Ontology" Web page: www.formalontology.it/Framing_first.htm.

[43] Quirk, R., Greenbaum, S., Leech, G., Svartvik, J., *A Comprehensive Grammar of the English Language*, Addison, Wesley, Longman, London, 1997.

[44] Ralston, A, Reilly, E, Hemmendinger, D., eds., *Encyclopedia of Computer Science*, Nature Publishing Group, London, 2000.

[45] Russell, S., Norvig, P., *Artificial Intelligence: A Modern Approach*, Prentice-Hall, NJ, 1995.

[46] Sebesta, R., *Concepts of Programming Languages*, Addison-Wesley, Menlo Park, CA, 1996.

[47] Sowa, J., "*Conceptual Graphs*" Web page: http://users.bestweb.net/~sowa/cg/index.htm

[48] Sowa, J., "Ontology" Web page: http://www.jfsowa.com/ontology/index.htm

[49] Sowa, J., "*Tutorial on Conceptual Graphs*" Web page: http://users.bestweb.net/~sowa/cg/tut.htm

[50] Sowa, J., *Principles of Semantic Networks: Explorations in the Representation of Knowledge*, Morgan Kaufmann, San Mateo, CA, 1991.

[51] Sowa, J., *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks/Cole, Pacific Grove, CA, 2000.

[52] Sproat, R., Shih, C., Gale, W., Chang, N., "A stochastic finite-state word-segmentation algorithm for Chinese" Computational Linguist 24:377-404, 1996.

[53] Stede, M., *Lexical Semantics and Knowledge Representation in Multilingual Text Generation*, Kluwer, Norwell, MA, 1999.

[54] Trujillo, A., *Translation Engines: Techniques for Machine Translation*, Springer, London, 1999.

[55] Tucker, A., ed., *Computer Science and Engineering Handbook*, CRC Press, Boca Raton, FL, 1997.

[56] van Benthem, J., ter Meulen, A., eds., *Handbook of Language and Logic*, MIT Press, Cambridge, MA, 1997.

[57] van Deemter, K., Peters, S., eds., *Semantic Ambiguity and Underspecification*, CSLI Publications, Stanford, CA, 1996.

[58] Wilson, W., *Essentials of Logic*, Research and Education Association, Piscataway, NJ, 1999.

[59] Wilss, W., *The Science of Translation: Problems and Methods*, Gunter Narr Verlag, Tübingen, 1982.

[60] Winograd, T., *Understanding Natural Language*, Academic Press, New York, 1972.

[61] Wittgenstein, L, *Philosophical Investigations*, Macmillan, London, 1953.

[62] Wong, S., "Machine Translation Techniques" Web page: http://www.fi.muni.cz/usr/wong/teaching/mt/notes/

[63] Zurada, J., *Introduction to Artificial Neural Systems*, West Publishing, St. Paul, MN, 1992.

… and many, many more, which were read in preparation for the comprehensive exam but did not find their way into this paper.