

Komunikace agentů

Komunikací se zde rozumí úmyslná výměna informací pomocí vytváření a příjmu znamení, odvozených ze sdíleného systému konvenčních znamení.

Lidé používají omezené množství konvenčních znamení (úsměv, potřásání rukou...) ve srovnání se zvířaty, na druhé straně vyvinuli komplexní, strukturovaný systém znamení, známý pod názvem **jazyk**, který jim umožňuje komunikovat o většině věcí, které znají o světě (prostředí, v němž se nacházejí). Přestože šimpanzi, delfini a jiní savci používají slovník se stovkami znamení a mají schopnost je řetězit, lidé jsou jediní tvorové, kteří spolehlivě komunikují pomocí neomezeného množství kvalitativně odlišných zpráv. Turingův test schopnosti umělé inteligence je založen právě na použití jazyka.

Komunikace jako akce

Agenti mohou mít jako jednu z dostupných akcí také schopnost produkovat jazyk. Tato činnost se nazývá **akce řeči (komunikace)**, přičemž pojem “řeč” má význam “volná řeč”, nikoliv jen “mluvení”, takže k akci řeči se také počítá např. psaní, znaková řeč, apod. Z tohoto hlediska také pojem **slovo** má širší význam.

Důvodem pro to, aby agent komunikoval při obvyklé činnosti je potřeba spolupráce s jinými agenty při řešení nějakého problému. Skupina agentů pak může kolektivně nebo individuálně získávat nějaké výhody, pokud může dělat např. něco z následujících činností:

- **Informovat** jeden druhého o části prozkoumaného světa, takže mají k dalšímu zkoumání menší část světa.
- **Dotazovat se** na konkrétní vlastnosti světa.
- **Odpovídat** na dotazy.
- **Žádat** (přikazovat) jiné agenty, aby vykonali nějakou akci.
- **Slibovat** provedení nějaké akce nebo **nabízet** podmínky.
- **Potvrzovat** žádosti a nabídky.
- **Sdílet** dojmy a zkušenosti (získané znalosti) s jinými.

Obtížné je pro agenta určit, *kdy* provést akci komunikace (aby komunikace byla opravdu zapotřebí) a *jaká* komunikační akce se má udělat. Z určitého hlediska se jedná o **plánovací problém**—agent si může vybrat ze souboru možných akcí a nějak musí otestovat, kterou akci vybrat k dosažení cíle (předání informace jinému agentovi).

Všechny plánovací potíže se projevují i při výběru komunikační akce. Není např. možné plánovat konverzaci na úrovni individuálních pohybů úst a jazyka, takže je nutné použít několik úrovní hierarchické abstrakce—přinejmenším slova, fráze a věty.

Dalším problémem je nedeterminismus: příkaz/žádost jednoho agenta může nebo nemusí být proveden/provedena, takže je zapotřebí mít k dispozici podmíněné plány. Konverzace se nedá naplánovat od začátku do konce, pouze je možné zkonstruovat obecný plán pro konverzaci, vygenerovat první větu, pak počkat na vjem odpovědi, a reagovat na odpověď.

Problém porozumění akce řeči je podobný jiným problémům porozumění (porozumění obrazovému vjemu, medicínské diagnóze, apod.). K dispozici je soubor nejednoznačných vstupů, z nichž je nutno zpětně vypracovat rozlišení stavu světa, který vstup vyprodukoval. Část problému porozumění je závislá na jazyku—je nutno znát syntaxi a sémantiku jazyka k pochopení důvodu, proč jiný agent vykonal konkrétní komunikační akci.

Problém porozumění zahrnuje i obecnější problém **rozpoznání plánu**. Pozorování agenta, který se otáčí a pohybuje směrem ke zlatu, může vést k vytvoření závěru, že agent má jako cíl získat zlato. Podobný druh vytváření mentálního modelu se vyžaduje také v situacích, kdy se určitý agent otočí směrem ke zlatu a řekne: “Seberu *to*.” I když v okolí budou také jiné objekty, mělo by být jasné, že “*to*” zde znamená “*zlato*”.

Část problému porozumění lze řešit pomocí logického usuzování (logické implikace jsou vhodným způsobem pro popis možností kombinovat slova a fráze na delší fráze). Jiná část problému se dá řešit technikami nejistého (přibližného) usuzování, např. fuzzy logikou aj. Obvykle existuje několik stavů světa, které všechny vedou k téže komunikační akci, takže příjemce zprávy musí rozhodnout, které stavy jsou pravděpodobnější.

Uvedené záležitosti jsou spojeny s tzv. vědou zabývající se přirozenými záležitostmi—např. porozumění přirozenému jazyku a jeho zpracování.

Podstata jazyka

V zásadě lze odlišovat **formální jazyky**—např. C++ a logika prvního řádu jsou rigidně definovány—a **přirozené jazyky**—např. čínština, angličtina, čeština, používané lidmi ke vzájemné komunikaci.

Umělá inteligence využívá pro formální jazyky notaci v tzv. Backus-Naurově formě (BNF). **Formální jazyk** je soubor řetězců složených ze sekvencí symbolů, které jsou převzaty z konečné množiny **terminálních symbolů**. Určitou potíží při práci s formálními a přirozenými jazyky je velké množství různých formalismů a notací pro zápis gramatik—většina je však založena na myšlence **frázové struktury**, tj. řetězce jsou složeny z podřetězců zvaných **fráze** z různých kategorií (podstatná jména – *noun phrases NP*, slovesa – *verb phrases VP*, atd.).

Fráze jsou vhodným prostředkem, jemuž lze přiřazovat významy (sémantika). Kategorizace frází napomáhá popsat přípustné řetězce jazyka. Lze říci, že jakákoliv fráze podst. jména *NP* může být kombinována (při dodržení určitých podmínek) se slovesnou frází *VP* pro vytvoření fráze z kategorie **věta** (*sentence S*).

Bez těchto pojmů a pravidel by bylo velmi obtížné (nebo i nemožné) vysvětlit, proč např. “v zásadě lze odlišovat formální jazyky a přirozené jazyky” je věta a např. “odlišovat v a jazyky jazyky formální lze přirozené” věta není. V angličtině např. “*the wumpus is dead*” je věta a “*wumpus the dead is*” věta není.

V různých jazycích existují ovšem různé gramatiky také vlivem odlišné podstaty (v angličtině je slovosled vázán mnohem přísnějšími pravidly než v češtině, takže přehození slov v češtině nemusí zrušit správnost kategorie věty, v angličtině však ano). Gramatické kategorie a pravidla jsou součástí jazykovědy.

Kategorie typu *NP*, *VP* a *S* jsou **neterminální symboly**. V notaci BNF se prepisovací pravidla skládají z jednoho neterminálního symbolu na levé straně a sekvence terminálů či neterminálů na pravé straně. Význam pravidla typu

$$S \rightarrow NP VP$$

je v tom, že dává možnost vzít jakoukoliv frází *NP*, připojit ji k jakékoliv frází *VP* a výsledkem je fráze typu *S*.

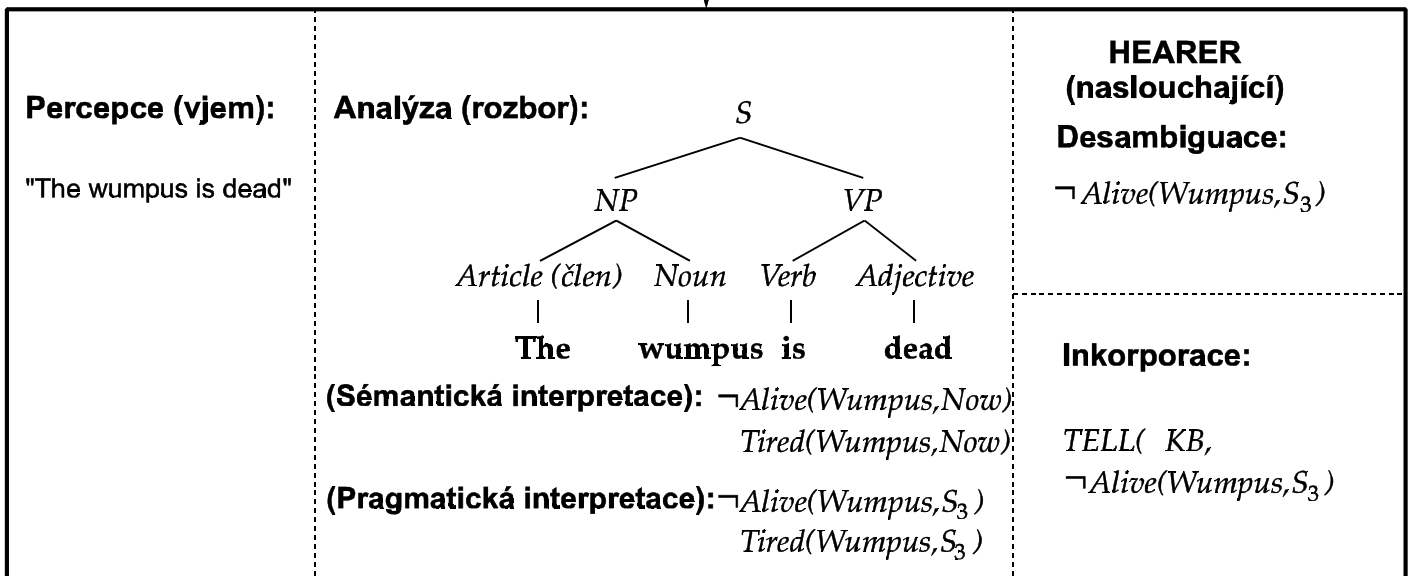
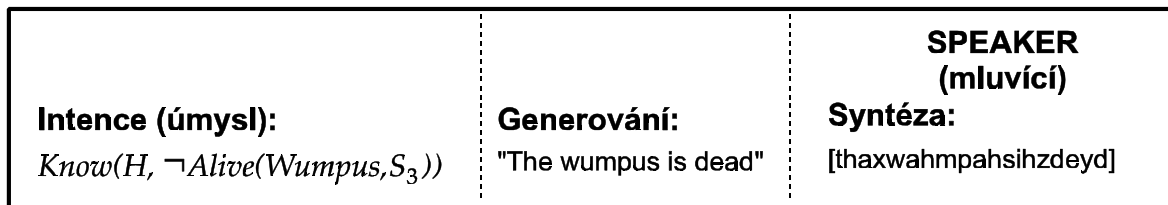
Komponenty komunikace

Typická komunikační epizoda, kdy řečník (speaker) S chce sdělit proposici P naslouchajícímu (hearer) H pomocí slov (words) W , se skládá ze sedmi procesů. Tři procesy souvisejí s řečníkem:

- Intence: S chce, aby H uvěřil P (typicky S sám věří P).
- Generace: S vybere slova W (protože vyjadřují význam P).
- Syntéza: S vysloví slova W (a obvykle je adresuje na H).

Čtyři procesy souvisejí s naslouchajícím:

- Percepce: H vnímá W' (ideálně $W' = W$, ale je možné nějaké zkreslení vjemu apod.).
- Analýza: H inferuje, že W' má možné významy P_1, \dots, P_n (slova a fráze mohou mít několik významů).
- Desambiguace: H inferuje, že S měl v úmyslu sdělit P_i (ideálně $P_i = P$, ale je možná chybná interpretace vjemu).
- Inkorporace: H se rozhodne uvěřit (přijmout) P_i (nebo P_i odmítne, pokud je mimo to, čemu již věří).



Předchozí obrázek ukazuje sedm procesů v kontextu příkladu (je použit příklad anglické fráze “*The wumpus is dead*”, přičemž “*wumpus*” byla kdysi jedna z prvních počítačových her, kde agent prohledává jeskyni skládající se z částí propojených chodbami—samotný wumpus byla bestie, skrývající se v některé části a sežrala každého, kdo k ní vlezl; navíc v některých částech byly bezedné šachty, do nichž bylo možno se propadnout, ovšem wumpus byl tak velký, že se do šachty nevešel, takže se nemohl propadnout; jediným zmírněním potíží byla občasná možnost nalézt hroudu zlata).

- Intence:** Z nějakého důvodu chce řečník sdělit něco potřebného naslouchajícímu—zde řečník *S* má důvod informovat *H*, že wumpus již dále není živ.
- Generace:** *S* použije znalost jazyka k rozhodnutí, co a jak říci o daném faktu úmrtí wumpuse. Výběr slov a jejich uspořádání je zatím větší problém než inverse, tj. rozumět sdělení, protože dosud bylo věnováno větší úsilí pochopit sdělení než vygenerovat sdělení ve srozumitelné formě.
- Syntéza:** Většina systémů umělé inteligence syntetizuje tištěný výstup, což je v podstatě triviální úloha. Syntéza řeči (aby zněla jako od člověka) nyní patří k populárním výzkumným úkolům. Na obrázku agent syntetizuje řeč pomocí řetězce zvuků přepsaného do jednoho z možných typů fonetické abecedy pro angličtinu: [*thaxwahmpahsihzdeyd*]. Detaily zde nejsou důležité; podstatné je, že syntetizované zvuky jsou něco odlišného od toho, co agent generuje. Fonetický řetězec je navíc nepřerušován mezerami, aby realizoval skutečnou rychlou řeč.
- Percepce:** Je-li médiem řeč, pak krok percepce se nazývá **rozpoznání řeči**; je-li médiem tisk, pak jde o **optické rozpoznání znaků**. Obě oblasti se již staly běžnými počítačovými záležitostmi. Zde předpokládáme perfektní schopnost *H* přijmout a dekodovat slova ze zvuku.
- Analýza:** Analýza zde má dvě hlavní části: *syntaktickou interpretaci* (rozbor) a *sémantickou interpretaci*. Sémantická interpretace zahrnuje jak porozumění významu slov, tak i pragmatickou interpretaci (znalost o současné situaci). *Rozbor* znamená proces přiřazování části jazyka (podstat-

né jméno, sloveso, atd.) každému ze slov ve větě a seskupením slov do frází. Jednou z možností je vytvoření **stromu rozboru** (viz předchozí obrázek). V tomto stromu vnitřní uzly reprezentují fráze, hrany použití gramatických pravidel, a listy představují slova.

Sémantická interpretace je proces získání významu výroku vyjádřeného v nějakém representačním jazyku. Na předchozím obrázku jsou dvě možné sémantické interpretace: *wumpus je mrtev* a *wumpus je unaven* (viz různé významy anglického slova *dead*). Výroky s několika možnými interpretacemi jsou tzv. **nejednoznačné**. Zde je použita logika pro reprezentaci, ale používají se i jiné reprezentace. **Pragmatická interpretace** je část sémantické interpretace, která bere do úvahy okamžitý stav—zde je konstanta *Now* (nyní) nahrazena jinou, S_3 , pro danou situaci.

Desambiguace: Většina řečníků nemusí být úmyslně nejednoznačná, ale mnoho výroků přirozeného jazyka má několik legálních interpretací. Komunikace funguje proto, že H odhaduje, která interpretace odpovídá sdělení přijímanému od S . Desambiguace je založena na pravděpodobnosti a přibližném usuzování (vliv nejistoty). Analýza vytváří možné interpretace; desambiguace vybírá tu nejlepší.

Inkorporace: Zcela naivní agent může věřit všemu, co je mu sdělováno, ale důmyslnější agent zachází se slovy W a odvozenou interpretací P_i jako s přídavnými částmi uvažovanými spolu s dalšími.

Předpokladem komunikace je, že a) agenti používají stejný jazyk, b) sdílejí stejný kontext a c) jsou alespoň částečně racionální (aby správně reagovali na vjem řeči).

Dva modely komunikace

K současným nejvýznamnějším směrům studia komunikace agentů patří způsoby změny reprezentace “vnitřních názorů” agentů na slova, a naopak změny přijímaných slov na “vnitřní názory” agentů, kde *vnitřní názor* znamená znalost uloženou v nějaké vhodné formě v agentově znalostní bázi (*znalostní báze* se zde opět rozumí soubor reprezentací faktů o světě—znalostní báze poskytuje odpovědi na dotazy např. přes dříve zmíněné funkce `Tell` a `Ask` v kapitole *Plánování*). Dva z existujících modelů jsou následující:

1. **Kódované zprávy:** řečník *S* má na mysli určitou proposici *P* a zakóduje ji do slov (nebo nějakých jiných příznaků) *W*. Naslouchající *H* se pokusí dekodovat zprávu *W* k získání původní proposice *P* (např. Morseův kód—z teček a čárek, používaný kdysi pro výměnu zpráv). V tomto modelu se v ideálním případě přenáší původní význam z hlavy *S* prostřednictvím přenášené zprávy a interpretace *H*, aniž by došlo k jakékoliv změně obsahu. Změnu obsahu však může způsobit např. šum komunikačního kanálu nebo chyba při kódování/dekódování.

Limitace modelu kódovaných a dekodovaných zpráv vedla k jinému modelu:

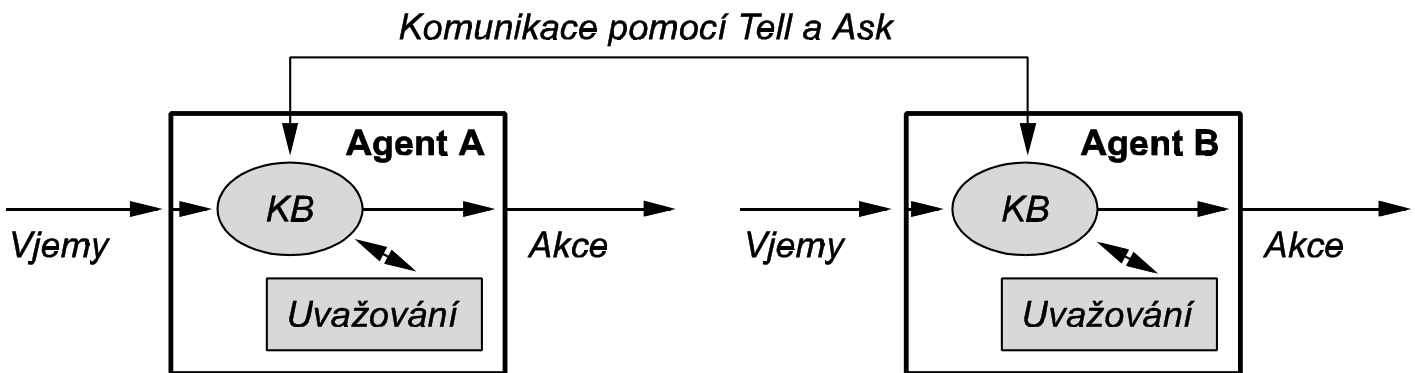
2. **Situační jazyk:** význam zprávy závisí jak na slovech, tak i na situaci v níž jsou slova vyslovena. Funkce kódování/dekódování zde mají navíc argument, který reprezentuje konkrétní, okamžitou situaci. Souvisí to se skutečností, že tatáž slova mohou mít v různých situacích zcela různé významy: “Já jsem teď zde” znamená pro Pepu v Praze v pondělí něco zcela jiného než pro Frantu v úterý v Ostravě, apod. Situační jazykový model má rovněž možný zdroj selhání komunikace: mají-li *S* i *H* odlišné představy o aktuální situaci, pak nemusí být zpráva předána na základě původního úmyslu.

Agenti mohou také komunikovat dvěma různými způsoby: oba sdílejí společný vnitřní representační jazyk, nepotřebují tedy žádný externí jazyk. Jiní agenti nemusí splňovat žádný předpoklad o vzájemném vnitřním jazyku, ale sdílejí komunikační jazyk (např. podmnožinu angličtiny).

Komunikace pomocí Tell a Ask

Při tomto typu komunikace agenti sdílejí tentýž vnitřní representační jazyk a mají přímý přístup ke svým znalostním bázím přes rozhraní Tell a Ask.

Agent A může předat proposici P agentovi B pomocí Tell (KB_B , " P ") stejně, jako by přidal P do své znalostní báze pomocí Tell (KB_A , " P "). Obdobně A může zjistit, zda B zná Q pomocí Ask (KB_B , " Q "). V umělé inteligenci, zabývající se uvedeným způsobem komunikace, se daný model nazývá **telepatické spojení**. Schématický diagram ukazuje modifikaci agentů (pro umožnění vzájemného přístupu ke svým znalostním bázím) přidáním vstup-výstupního zařízení ke KB , tj. navíc k potřebným zařízením pro vstup/výstup vjemů a akcí:



Lidé nemají schopnosti telepatické komunikace, takže uvedený způsob komunikace nemohou používat, ale je zcela možné naprogramovat skupinu robotů pomocí společné interní reprezentace a vybavit je rádiovými nebo infračervenými propojeními pro přímý přenos vnitřní reprezentace znalostí mezi jejich znalostními bázemi přes Tell a Ask.

V takovém případě je nutné, aby agenti kromě stejného vnitřního formátu reprezentace jazyka sdíleli mnoho symbolů. Některé symboly jsou *statické*: lze předem pevně stanovit význam (konstanty, popis neproměnné části světa, apod.). Jiné symboly jsou *dynamické*: vytvářejí je agenti tehdy, když začnou prozkoumávat svět.

Se statickými symboly potíže nevznikají, ale obtížné bývá *synchronizovat* použití dynamických symbolů.

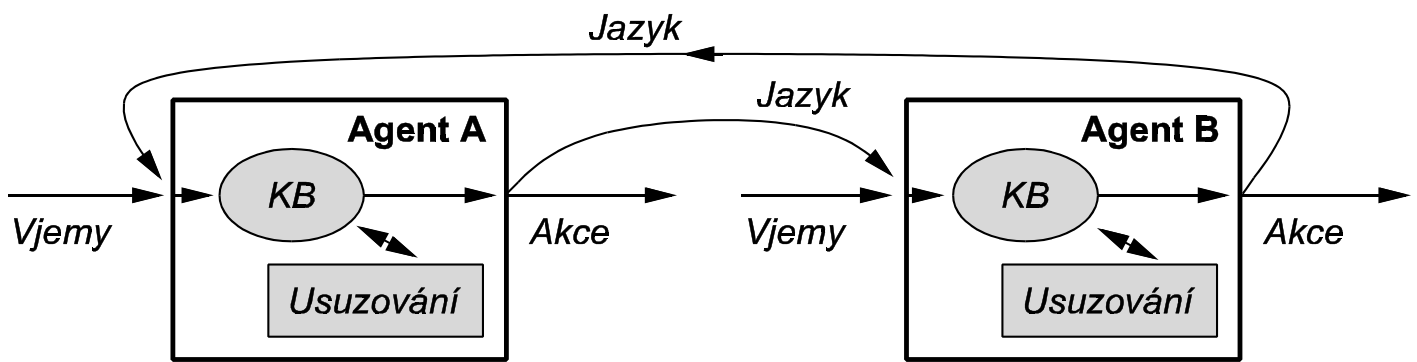
Tři potíže se synchronizací dynamických symbolů:

1. Musí existovat takový postup přidělování názvů, aby A i B simultánně nezaváděli tentýž symbol pro věci s různým významem. Jednou možností je zavedení agentova jména jako indexu nově zaváděného symbolu, např. " $At(Home_A)$ " znamená "*doma u agenta A*".
2. Je zapotřebí, aby existoval způsob stanovení relace pro symboly zavedené různými agenty, např. aby určitý agent byl schopen říci, zda P_{A1} a třeba P_{B2} znamená stejné místo nebo ne. Problém narůstá s počtem agentů.
3. Není snadné uvést v soulad rozdíly mezi znalostními bázemi různých agentů. Pokud je komunikace volná a bezprostřední, pak všichni agenti mohou postupovat tak, že rozešlou sdělení o nové skutečnosti všem ostatním hned, jak se skutečnost dozví, takže každý bude mít stejnou znalost. V realitě je však taková komunikace omezená, např. šířkou pásma přenosu nebo tím, že nejsou neustále v propojení. Když se dostanou zpět do kontaktu, nastává problém s určením toho, co z nové informace stojí za komunikaci a která zajímavá fakta ostatní agenti znají.

Existuje přirozeně mnoho dalších problémů, např. telepatičtí agenti, jak byli popsáni, jsou zranitelní sabotáží (nějaký protřelý agent použije `Te11` k šíření lží a naivní agent všemu uvěří).

Komunikace pomocí formálního jazyka

Z důvodů možnosti sabotáže a/nebo nemožnosti používat tentýž vnitřní jazyk komunikuje většina agentů pomocí jazyka než prostřednictvím přímého přístupu do znalostní báze. Následující diagram ukazuje typ komunikace pomocí jazyka: agenti provádějí akce, které produkují jazyk, který mohou ostatní agenti přijímat (vnímat). Externí komunikační jazyk se může samozřejmě lišit od interního representačního jazyka, a každý agent může mít rozdílný vnitřní jazyk. Nemusí se shodnout na žádném interním symbolu, pokud každý z agentů je schopen spolehlivě mapovat z externího jazyka do svých interních symbolů.



Externí komunikační jazyk s sebou přináší problémy spojené s generováním a analýzou, a řada výsledků výzkumu z oblasti zpracování přirozeného jazyka (NLP) se zde využívá. *Nejobtížnější* na komunikaci pomocí jazyka je stále problém č. 3 z předchozí podkapitoly: uvést rozdíly ve znalostních bázích agentů v soulad. To, co říká agent *A* a jak agent *B* interpretuje výrok od *A*, závisí velmi silně na vnitřních “názorech” obou agentů (včetně toho, jaký mají “názor” na “názor” toho druhého agenta). Znamená to, že agenti mající tentýž interní a externí jazyk nemusí mít potíže s generováním a analýzou sdělení, ale stále pro ně může být obtížné rozhodnout, co si navzájem sdělit.

V dalších úvahách se předpokládá využití omezené podmnožiny angličtiny pro komunikaci agentů pomocí externího jazyka.

Formální gramatika pro podmnožinu angličtiny

Nechť je pro komunikaci použita omezená podmnožina angličtiny ϵ_0 (např. pro komunikaci agentů ve zmíněném světě *wumpus*). Komunikační jazyk lze do značné míry popsat pomocí technik pro formální jazyky, např. využití pevné množiny písmen (pro psaný jazyk) nebo zvuků (pro mluvený jazyk). Symboly jsou formovány do řetězců, řetězce do frází apod. Ne celý jazyk lze takto popsat, protože např. není jasné, co vše do jazyka patří (existují dialekty, rozdíly např. v kanadské, britské, australské a americké angličtině, vznikají přirozené časové změny—angličtina ze středověku je odlišná od současné, přičemž obdobné záležitosti platí i pro jiné jazyky, definice lingvistů se mohou lišit, apod.). Kromě toho existuje mnoho ngramatických výroků, kterým lze bez problémů porozumět.

Řada frází a vět nemusí být správných na vyšší gramatické úrovni, ale z části jsou správné, a pro různé mluvčí se mohou pohybovat od správného k nesprávnému; příklady z angličtiny:

To whom did you send the letter?

Next to whom did you stand?

Of whom did you meet a friend?

Of whom did you see the dealer that bought the picture that Vincent painted?

Skutečně obtížnou částí zpracování přirozeného jazyka je *sémantická interpretace* (stanovení významu) a *desambiguace* (odstranění mnohovýznamnosti). V programovacích jazycích je v principu každé vyjádření příkazem, zatímco v přirozeném jazyce musí naslouchající rozpoznat, zda jde o příkaz, dotaz, sdělení, slib... Dále bude jako příklad použita formální definice jazyka ϵ_0 .

Lexikon ϵ_0

Lexikon je seznam povolených slov, která jsou seskupována do kategorií či částí řeči srozumitelné uživateli slovníku: *podst. jména*, *zájmena*, *názvy* označují věci, *slovesa* označují události, *příd. jména* modifikují podstatná jména, a *příslovce* modifikují slovesa. Dalšími kategoriemi jsou *členy* (a, an, the), *předložky* (in, on, at, ...), a *spojky* (and, or, ...). Příklad ukazuje malý lexikon pro svět hry *wumpus*:

podstatné jméno → *stench* | *breeze* | *glitter* | *nothing* | *wumpus* | *pit* | *pits* | *gold* | *east* | ...

sloveso → *is* | *see* | *smell* | *shoot* | *feel* | *stinks* | *go* | *grab* | *carry* | *kill* | *turn* | ...

přídavné jméno → *right* | *left* | *east* | *south* | *back* | *smelly* | ...

příslovce → *here* | *there* | *nearby* | *ahead* | *right* | *left* | *east* | *south* | *back* | ...

zájmeno → *me* | *you* | *I* | *it* | ...

název → *John* | *Mary* | *Boston* | *Aristotle* | ...

člen → *the* | *a* | *an* | ...

předložka → *to* | *in* | *on* | *near* | ...

spojka → *and* | *or* | *but* | ...

číslice → **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9**

Každá kategorie v zobrazeném příkladu má na konci tři tečky ... k tomu, aby bylo zjevné, že existují i další slova v kategorii. Pro chybějící slova jsou dva důvody:

1) *podst. jména, slovesa, příd. jména a příslovce*: nelze je vyjmenovat všechny, protože je jich veliké množství v každé třídě a navíc jich neustále přibývá (např. slovo *fax* je běžně používáno teprve omezenou dobu); tyto čtyři kategorie tvoří tzv. **otevřenou třídu**.

2) Ostatní uvedené kategorie tvoří tzv. **uzavřenou třídu**: mají malý počet slov, které v principu lze vyjmenovat a mění se spíše během staletí než měsíců (např. zájmena *thee* a *thou* byla obecně užívána v 17. století, v 19. byla na ústupu a v současnosti je lze nalézt jen v poezii a některých regionálních dialektech).

Gramatika ϵ_0

Slova je nutno kombinovat do frází. V ϵ_0 se použije pět neterminálních symbolů k definici různých druhů frází: *věta S (sentence)*, *fráze podst. jména NP (noun phrase)*, *fráze slovesná VP (verb phrase)*, *předložková fráze PP (prepositional phrase)*, a *vztažná věta RelClause (relative clause)*—následuje a modifikuje NP, např. *The agent **who gave me the gold** is there*. Například pro svět *wumpus*:

<i>S</i>	→	<i>NP VP</i>	I + feel a breeze
		<i>S spojka S</i>	I feel a breeze + and + I smell a wumpus
<i>NP</i>	→	<i>zájmeno</i>	I
		<i>podst. jméno</i>	pits
		<i>člen podst. jméno</i>	the + wumpus
		<i>číslice číslice</i>	3 4
		<i>NP PP</i>	the wumpus + to the east
		<i>NP RelClause</i>	the wumpus + that is smelly
<i>VP</i>	→	<i>sloveso</i>	stinks
		<i>VP NP</i>	feel + a breeze
		<i>VP příd. jméno</i>	is + smelly
		<i>VP PP</i>	turn + to the east
		<i>VP příslovce</i>	go + ahead
<i>PP</i>	→	<i>předložka NP</i>	to + the east
<i>RelClause</i>	→	<i>that VP</i>	that + is smelly

Syntaktická analýza (rozbor)

Existuje mnoho algoritmů pro syntaktickou analýzu. Zde je uveden velmi jednoduchý algoritmus, který nedeterministicky vybírá jeden možný **strom rozboru** (parse tree), pokud existuje. Zpracovává seznam slov jako **les rozborů** (parse forest): uspořádaný seznam stromů rozboru. V každém kroku hlavní smyčky nalézá nějakou posloupnost elementů lesa, která odpovídá pravé straně některého gramatického pravidla. Pak nahradí posloupnost jednoduchým stromem rozboru, jehož kategorie je na levé straně pravidla a jehož potomci jsou uzly v původní posloupnosti:

```

function BOTTOM-UP-PARSE(words, grammar) returns a parsetree

  forest ← words
  loop do
    if LENGTH(forest)= 1 and CATEGORY(forest[1]) = START(grammar) then
      return forest[1]
    else
      i ← choose from {1 ... LENGTH(forest)}
      rule ← choose from RULES(grammar)
      n ← LENGTH(RULE-RHS (rule))
      subsequence ← SUBSEQUENCE(forest, i, i+n-1)
      if MATCH(subsequence, RULE-RHS(rule)) then
        forest[i...i+n-1] ← [MAKE-NODE(RULE-LHS(rule), subsequence)]
      else fail
  end

```

Pro ukázkou je uvedeno, jak by rozbor probíhal pro řetězec “*the wumpus is dead*”:

<i>forest</i>	<i>subsequence</i>	<i>rule</i>
The wumpus is dead	The	Article → <i>the</i>
Article wumpus is dead	wumpus	Noun → <i>wumpus</i>
Article Noun is dead	Article Noun	NP → Article Noun
NP is dead	is	Verb → <i>is</i>
NP Verb dead	dead	Adjective → <i>dead</i>
NP Verb Adjective	Verb	VP → Verb
NP VP Adjective	VP Adjective	VP → VP Adjective
NP VP	NP VP	S → NP VP
S		

(Detailní informace o využívání gramatiky poskytuje oblast NLP.)

Sémantická interpretace

Sémantika formálních jazyků je na rozdíl od jazyků přirozených poměrně snadná. Není obtížné zjistit význam zápisů “ $X+Y$ ” v aritmetice nebo “ $X \wedge Y$ ” v logice, protože mají tzv. **kompoziční sémantiku**: sémantika libovolné fráze je funkcí sémantiky sub-frází; nezáleží na jakékoliv jiné frázi před, po nebo zahrnuté v dané frázi. Je-li znám význam X a Y (a také +), pak je jasný význam celé fráze.

Mimo jednoduchý jazyk aritmetiky nebo logiky je situace mnohem složitější. Např. v nějakém programovacím jazyku může mít výraz “10+10” sémantickou interpretaci 4, pokud se výraz vyskytne v delším řetězci:

```
Base(2);  
x ← 10+10;
```

Ovšem i zde lze obhájit kompoziční sémantiku, pokud se vyjde z tvrzení, že sémantická funkce asociovaná s “ $x+y$ ” je sečíst x a y pro daný číselný základ.

Velkou výhodou kompoziční sémantiky je totéž jako u tzv. bezkontextových gramatik: umožňuje zpracovávat nekonečnou gramatiku pomocí konečné (a často malé) množiny pravidel.

Z povrchního pohledu by se zdálo, že přirozené jazyky nemají kompoziční sémantiku. Ve frázi “*The batter hit the ball*” (pálkař trefil míček) lze očekávat sémantickou interpretaci slova “*batter*” jako “*někdo, kdo máchá pálkou*” a “*ball*” jako “*sférický sportovní předmět*”. Ale ve frázi “*The chef mixed the batter to be served at the ball*” (kuchař namíchal těsto tak, aby se podávalo jako kuličky) se dá očekávat, že dvě slova mají jiný význam. Z toho by se dalo usoudit, že “*ball*” a “*batter*” závisí nekompozičně na okolním kontextu.

Tyto sémantické interpretace jsou však pouze předpokládané a preferované, nikoliv jediné možné. Slovo “*ball*” také znamená “ples, bál”. Je zcela *možné*, že “*The batter hit the ball*” se vztahuje k tomu, že směs koláčů vypadala velkolepě na formálním tanci—při dostatečném úsilí lze vytvořit příběh, kde uvedený význam je preferovaný.

Shrnutí: *Samotná sémantická interpretace nemůže zajistit správnou interpretaci fráze nebo věty. Kompoziční přístup je v moderních systémech zabývajících se přirozenými jazyky jeden z nejpoužívanějších.*

Je ale zapotřebí použít dva kroky: **sémantická interpretace** odpovídá za kombinaci významů kompozičně tak, aby byly získány možné interpretace, a **desambiguace** (odstranění nejednoznačnosti) odpovídá za výběr nejlepší interpretace.

Sémantika výrazu “John loves Mary”

Následující popis se týká vytvoření gramatiky pro velmi malou podmnožinu angličtiny. Prvním krokem je stanovení faktů—jaké sémantické reprezentace s kterými frázemi mají být asociovány. Bude použita jednoduchá věta “John loves Mary” asociována se sémantickou interpretací $Loves(John, Mary)$. Ve větě je snadné vidět, z kterých slov pocházejí které části sémantické interpretace.

Komplikovanější částí je rozhodnutí, jak ty části dát dohromady, zejména pro mezifráze typu VP “*loves Mary*”, kde VP je slovesná fráze. Sémantická interpretace dané fráze není ani logický výraz, ani kompletní logická věta. Intuitivně, “*loves Mary*” je popis, který lze nebo nelze aplikovat na konkrétní osobu (v tomto případě se vztahuje k osobě *John*). Znamená to, že “*loves Mary*” je **predikát**, který v kombinaci s termem reprezentujícím osobu (která miluje) dává kompletní logický výraz. Za použití λ -notace lze reprezentovat “*loves Mary*” jako predikát:

$$\lambda x Loves(x, Mary).$$

(λ -operátor se používá k zápisu funkce z výrazu, např. výraz $x^2 - y^2$ se přepíše na $\lambda x, y x^2 - y^2$, např. $(\lambda x, y x^2 - y^2)(25, 24) = 25^2 - 24^2 = 49$; funkce popisuje, co jsou argumenty a že se má udělat rozdíl druhé mocniny prvního a druhého argumentu; viz případně i jazyk *Lisp*.)

NP “*Mary*” lze reprezentovat logickou konstantou $Mary$. To vede k definici pravidla “ NP se sémantikou obj následováno VP se sémantikou rel dává větu S , jejíž sémantika je výsledkem aplikace relace rel na objekt obj ”:

$$S(rel(obj)) \rightarrow NP(obj) VP(rel)$$

Pravidlo říká, že sémantická interpretace “John loves Mary” je

$$(\lambda x Loves(x, Mary))(John)$$

což je ekvivalentní $Loves(John, Mary)$.

Protože slovesné fráze *VP* jsou representovány jako predikáty, je vhodné konsistentně representovat rovněž slovesa jako predikáty. Slovo “loves” je representováno jako

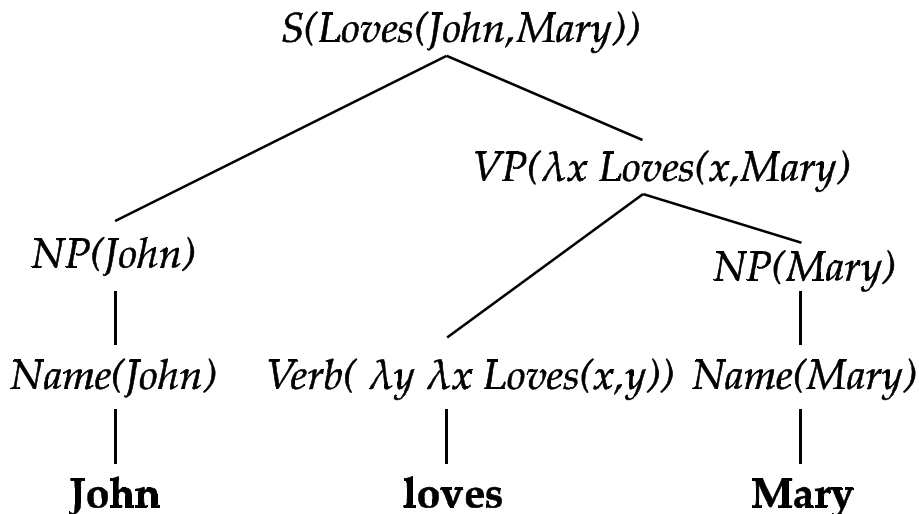
$$\lambda y \lambda x \text{ Loves}(x,y),$$

což je predikát, který s argumentem $y = \text{Mary}$ vrací predikát $\lambda x \text{ Loves}(x, \text{Mary})$.

Pravidlo $VP \rightarrow \text{Verb NP}$ (*verb phrase*=slovesná fráze, *verb*=sloveso) aplikuje predikát [který je sémantickou interpretací slovesa na objekt, který je sémantickou interpretací *NP* (noun phrase=podstatné jméno)] k získání sémantické interpretace celé slovesné fráze *VP*. Gramatika a strom rozboru (se sémantickými interpretacemi pro řetězec “John loves Mary”) jsou následující:

$$\begin{aligned} S(\text{rel}(\text{obj})) &\rightarrow NP(\text{obj}) VP(\text{rel}) \\ VP(\text{rel}(\text{obj})) &\rightarrow \text{Verb}(\text{rel}) NP(\text{obj}) \\ NP(\text{obj}) &\rightarrow \text{Name}(\text{obj}) \end{aligned}$$

$$\begin{aligned} \text{Name}(\text{John}) &\rightarrow \mathbf{John} \\ \text{Name}(\text{Mary}) &\rightarrow \mathbf{Mary} \\ \text{Verb}(\lambda x \lambda y \text{ Loves}(x,y)) &\rightarrow \mathbf{loves} \end{aligned}$$



Nejednoznačnost a její odstraňování

Nejednoznačnost (ambiguita) a její odstraňování (desambiguace) hraje výraznou roli v procesu komunikace (interpretační fáze). V ideálním případě má řečník *S* na mysli sdělení *P* a provede akci mluvení, která však může mít několik interpretací, a je otázkou, kterou z interpretací lze v dané situaci považovat za nejlepší vzhledem k předávání výroku *P*. Naslouchající *H* si je možné nejednoznačnosti vědom a *P* interpretuje—odstraňuje víceznačnost. Někdy může *H* být zmaten a požádat o vyjasnění, to ale při častém opakování vede ke zdržování komunikace, zejména pokud dojde k dalšímu, vícenásobnému “vyjasňování vyjasňovaného”. Existuje velmi mnoho narušení plynulé komunikace, ale ukazuje se, že *největším problémem je to, že většina toho, co je vyřčeno, je nejednoznačné*. Typickým příkladem jsou např. titulky zpráv (známým příkladem je novinový titulek z doby druhé světové války: *Americans pushes bottle up Germans*: i z hlediska překladu nelze *přesně říci*, co je významem—např. *bottle* může znamenat *láhev, odvaha, otýpka*, a nahromadění nejednoznačných pojmů do fráze činí interpretaci často nemožnou).

Lexikální ambiguita je nejjednodušším typem ambiguity, kde nějaké jedno slovo má více významů. Např. anglické přídavné jméno *hot* znamená *horký, kořeněný, nejnovější, nesnesitelný, dychtivý, rozčilený, skvělý, čerstvý, radioaktivní, sexy, populární, ukradený* aj. (v kterémkoliv přirozeném jazyce lze nalézt obdobné příklady). Lexikální ambiguita zahrnuje také nejednoznačnost kategorie slova: *back* je příslovce v *go back*, přídavné jméno v *back door*, podstatné jméno v *the back of the room*, a sloveso v *back up your files*.

Syntaktická ambiguita (strukturální ambiguita) se může vyskytnout jak s lexikální nejednoznačností, tak i bez ní. Např. *I smelled a wumpus in 2,2* (kde *2,2* znamená označení koordinát výskytu wumpuse v dané hře) má dva výsledky rozboru: v jednom propoziční fráze modifikuje podstatné jméno, v druhém modifikuje sloveso. Syntaktická ambiguita vede k další ambiguitě, k sémantické.

Sémantická ambiguita v uvedeném příkladu znamená, že buď wumpus je v *2,2* a je odtud cítit nebo zápach je cítit v *2,2* a wumpus je někde jinde (*cítit jsem (wumpuse v 2,2)* nebo (*v 2,2 jsem cítil*) wumpuse). Zde by chybná interpretace mohla být osudovou chybou vzhledem k zaměření wumpuse. Rovněž může dojít k tomu, že se sémantická nejednoznačnost objeví i ve frázích, kde není lexikální nebo syntaktická ambiguita. Např. fráze *coast road (pobřežní silnice)* může znamenat *silnici vedoucí po pobřeží* nebo *silnici vedoucí k pobřeží*.

Referenční ambiguita je všudypřítomná forma nejednoznačnosti. Anaforické (odkazové na předchozí kontext) výrazy typu *it* se mohou odkazovat téměř na vše. Referenční ambiguita se vyskytuje v komunikaci proto, že přirozené jazyky obsahují v silně převažující míře slova pro kategorie, nikoliv pro individuální objekty [neexistuje slovo pro *the-apple-I-had-for-lunch-today* (jablko, které jsem měl dnes k obědu), pouze kategorie typu *apple* (jablko)].

Pragmatická ambiguita se objevuje tehdy, když promlouvající *S* a naslouchající *H* se neshodnou na tom, co je okamžitá situace. *S* říká “*setkám se s tebou příští čtvrtek*” za předpokladu, že se mluví o 12. dni měsíce, ale *H* si myslí, že se jedná o 19., takže dojde k chybné komunikaci. Jiný příklad je výše v podkapitole *Dva modely komunikace* ve zmínce o situačním jazyku (slovo “zde”).

Lokální ambiguita se někdy objevuje ve frázi či větě, kde lze udělat rozbor podřetězce několika způsoby, ale pouze jeden z nich se shoduje s širším kontextem celého řetězce. Např. v programovacím jazyce *C* je význam řetězce “**c*” buď “*pointer na c*”, pokud se řetězec vyskytuje v deklaraci *char *c*; nebo “*vynásob hodnotou c*” pokud se vyskytuje ve výrazu “*2*c*”. Podobně v angličtině je podřetězec “*the radio broadcasts*” frází *NP* v širším kontextu “*the radio broadcasts inform*”, zatímco v jiném kontextu “*the radio broadcasts information*” jde o frázi *NP* “*the radio*” následovanou další frází *VP* “*broadcasts information*” (*broadcasts* je vysílání jako podstatné jméno v množném čísle nebo *vysílá* jako sloveso, tj. *rozhlasová vysílání informují* nebo *rozhlas vysílá informace*).

Fráze či věta může být syntakticky nejednoznačná, ale sémanticky jednoznačná. Např. “*S₁ a S₂ a S₃*” má dva různé rozborů “*(S₁ a S₂) a S₃*” a “*S₁ a (S₂ a S₃)*”, ale stejný význam vzhledem k asociativnosti konjunkce *a* (to ale nemusí být obecně platné a může záležet na typu jazyka a gramatiky).

Vágnost (ve smyslu nejasnost, neurčitost, nejistota) je v přirozených jazycích obvyklá. Výrok “*Venku je horko*” se týká teploty, ale termín “*horko*” je vágní, takže je možná vcelku široká interpretace (pro někoho je venku horko při 25°C, pro jiného při 32°C, apod.).

Nejednoznačnost může vyplynout i z typu akce řeči: Na dotaz “*Viš, kolik je hodin?*” odpoví *H* “*ano*” nebo může odpovědět hodnotou času, např. “*půl sedmé a tři minuty*”. Obojí může být pro tentýž dotaz správné, ovšem v závislosti na typu otázky (a např. kontextu, apod.).

Desambiguace

Desambiguace je v principu problém diagnózy, tj. rozboru a rozpoznání. Naslouchající H si udržuje model světa a na základě vyslechnutí nové akce řeči přidá možné interpretace toho, co slyšel, k modelu jako hypotézy. K interpretacím lze využít různé techniky zpracování nejistoty a přibližnosti (logika monotonická a nemonotonická, pravděpodobnosti, shlukování, stochastická simulace, Bayesovy metody, tvorba a zpracování pravidel, Dempster-Shaferova metoda, znalostní inženýrství, fuzzy usuzování, aj.).

K úspěšnému výsledku je zapotřebí, aby svět obsahoval co nejvíce informací o světě. Např. ke korektnímu vyřešení syntaktické ambiguity v anglické větě “*Joseph saw the Grand Canyon flying to New York*” (Josef viděl Grand Canyon při letu do New Yorku) je nutno vědět, že je mnohem pravděpodobnější, že Josef—a nikoliv Grand Canyon—letěl letadlem (a tedy že Josef neviděl Grand Canyon letící do New Yorku).

Je také zapotřebí mít dobrý model o názorech S a H , aby se dalo určit, co měl S v úmyslu říci. Např. obvyklá interpretace anglického sdělení “*I am not a crook*” je, že “*nejsem zloděj*”, i když lze sdělení také přeložit jako “*nejsem pastýřská berla*”. Podobně slovo “*bank*” znamená mj. “*břeh, lavice, banka*”, ale je značně pravděpodobné, o co jde ve výroku “*Jimmi keeps his money in the bank*”, tj. Jimmi asi nezahrabává své peníze do břehu (i když i to je možné).

Desambiguace obecně vyžaduje kombinaci čtyř modelů:

1. **Model světa:** pravděpodobnost výskytu faktu ve světě.
2. **Mentální model:** pravděpodobnost, že řečník S má úmysl sdělit fakt naslouchajícímu H , a to za předpokladu, že k výskytu faktu došlo (je vhodné uvážit i případy, kdy k faktu nedošlo a S o tom nemá správnou informaci nebo úmyslně lže). Zde se rovněž kombinuje model názorů S s modelem toho, jaký má S názor na názory H , atd.
3. **Model jazyka:** pravděpodobnost toho, že bude zvolen určitý řetězec slov za předpokladu, že S má intenci komunikace o určitém faktu.
4. **Akustický model:** pravděpodobnost vygenerování konkrétní sekvence zvuků za předpokladu, že S vybral daný řetězec slov (problém obecně souvisí s řešením úloh v oblasti příjmu a vysílání signálů—zpracování vizuálních a sluchových vjemů, jejich zpracování, vytváření a předávání—to je mimo rozsah předmětu).

Zcela jistým a posledním důvodem, proč správná interpretace je tak velmi obtížný problém, je skutečnost, že může existovat několik správných interpretací. Např. vtipy jsou často založeny na tom, že *H* se pobaví pomocí dvou současně platných interpretací. Většina programů, schopných porozumění, zatím ignoruje tyto skutečnosti (možnost současného výskytu několika případů).

Bezkontextové gramatiky neposkytují příliš užitečný jazykový model, protože gramatika neříká, který řetězec je pravděpodobnější než jiný—prostě dělí řetězce do dvou tříd: gramatické a negramatické.

Nejjednodušší formou je poskytnout pravděpodobnostní rozdělení pomocí **pravděpodobnostní bezkontextové gramatiky** (tzv. PCFG—probabilistic context-free grammar). V jazykovém modelu PCFG má každé prepisovací pravidlo asociovanou pravděpodobnost, a součet pravděpodobností pro všechna pravidla, mající tutéž levou stranu, je 1. Např.:

$$S \rightarrow NP VP \quad (0.9)$$

$$S \rightarrow S \text{ konjunkce } S \quad (0.1).$$

V modelu PCFG je pravděpodobnost řetězce, $P(\textit{words})$, pouhou sumou pravděpodobností jeho stromů rozboru (*parsing trees*)—jeden strom pro jednoznačný řetězec, žádné stromy pro negramatický řetězec a několik stromů pro nejednoznačný řetězec. Pravděpodobnost daného stromu je součinem pravděpodobností všech pravidel, která tvoří uzly stromu.

Problémem PCFG je bezkontextovost, neboli rozdíl mezi např. $P(\textit{“I ate a banana”})$ a $P(\textit{“I ate a bandana”})$ [tj. *“jedl jsem banán”* a *“jedl jsem šátek”* (bandana je typ barevného bavlněného nebo hedvábného šátku)], záleží výhradně na $P(\textit{“banana”})$ versus $P(\textit{“bandana”})$, a nikoliv na relaci mezi slovesem *“ate”* a odpovídajícími podstatnými jmény. K získání takového vztahu je zapotřebí nějaký druh kontextově závislého modelu.

Pozn.: V současnosti zřejmě převažují **pravděpodobnostní techniky** desambiguace, které vycházejí z existence tzv. textových **korpusů**, z nichž se odvozují příslušné statistiky a pravděpodobnosti. Existují i korpusy pro český jazyk. Samozřejmě záleží na tom, kdy a z čeho je korpus sestavován a jak je rozsáhlý, takže různé korpusy mohou vést k více či méně odlišným výsledkům.