

БАГАТОЗНАЧНА СМИСЛОВА ІНТЕРПРЕТАЦІЯ УСНОМОВНОГО СИГНАЛУ

Тарас ВІНЦЮК

Міжнародний науково-навчальний центр інформаційних технологій та систем НАН і Міністерства освіти

40, просп. Академіка Глушкова, Київ 252022

Тел.: +380 44 266-4356 Факс: +380 44 266-1570

Електронна пошта: vintsiuk@uasoiro.freenet.kiev.ua

Abstract

Significant continuous speech understanding algorithm is proposed. It consists in that for current discrete time step it is being found not only the $N \gg 1$ best initial word subsequences, which are permissible in the speech dialogue language, but at the next discrete time steps it is being examined only such words that could continue these $N \gg 1$ already accumulated permissible in the speech dialogue language initial word subsequences. By this way it is being removed lacks and united advantages of two earlier mentioned speech understanding models [1—4].

1. Вступ

В [1] запропоновано дві технології (моделі) машини для усного перекладу та/або диктування (МУПД). Вони є ієрархічно організованими. Перша модель базується на так званій генеративній моделі розпізнавання, розуміння та синтезу усномовного сигналу. При цьому синтез мовлення використовується як зворотній зв'язок в процесі розпізнавання. Три основні блоки МУПД такі: 1) модель зовнішнього світу, яка описує всі можливі смисли, що передаються в процесі діалогу, 2) генератор текстів або речень природної мови і 3) генератор фонетично-акустичних прототипів (модельних сигналів) усної мови. Проблема МУПД сформульована як 1) проблема знаходження для сигналу, що розпізнається, найбільш схожого на нього модельного сигналу злиганої мови з множини всіх можливих сигналів прототипів, породжених третім блоком, для всіх можливих текстів і речень природної мови, що генеруються другим блоком для всіх можливих смислів, що передаються першим блоком, а також як 2) проблема аналізу (розбору) останнього як послідовності слів та канонічної форми смислу, що передаються мовним сигналом. Отримані послідовності слів є граматично і семантично правильними, тому можуть бути надруковані диктувальною машиною або синтезовані каналом третього блоку для іншої мови в машині для усного перекладу.

Другий шлях створення МУПД — так звана багаторівнева модель з багатозначними рішеннями. Тут не використовується синтез мовлення в якості зворотнього зв'язку. Натомість вводяться

багатозначні рішення за спрощених умов на всіх рівнях ієрархії оброблення мовного сигналу. Наприклад, на першому рівні вирішується узагальнена проблема розпізнавання злигано мовлення, яка полягає в тому, що виходячи з припущення про вільний порядок слів знаходяться $N1 \gg 1$ найкращих послідовностей слів. Потім на другому рівні ці багатозначні послідовності аналізуються, допоки не знайдуться послідовності слів, що збігаються з породженими генератором текстів або речень природної мови. Таким чином, на другому, найвищому рівні отримуємо $N2 \gg 1$ найкращих результатів розуміння, з них остаточно обирається один найкращий.

Далі пропонується алгоритм багатозначної смислової інтерпретації усної мови для МУПД. Він усуває недоліки та поєднує переваги двох вищезгаданих моделей.

Цей алгоритм полягає в тому, що, наприклад, у дворівневій багатозначній МУПД на першому рівні вирішується узагальнена проблема розпізнавання не за умов припущення про вільний порядок слів, але в умовах розгляду $N \gg 1$ найкращих послідовностей слів, або точніше $N \gg 1$ початкових підпослідовностей слів, що мають бути допустимими в мові усного діалогу, тобто відповідати синтаксису, семантиці та прагматиці усного діалогу. Тому для поточного відліку часу будемо знаходити лише $N \gg 1$ найкращих початкових послідовностей слів, котрі допустимі в мові усного діалогу, а на наступному часовому кроці розглядатимемо тільки ті слова, які можуть продовжувати вже відібрані підпослідовності слів, які є допустимими в мові усного діалогу. Таким чином, вводимо багатозначну смислову інтерпретацію усномовного сигналу.

В результаті досягається така ж точність розуміння, як і в другій моделі, але за значно менших обсягів обчислень та пам'яті.

Далі в деталях описується основний алгоритм багатозначної смислової інтерпретації.

2. Специфікації природної мови

Описуватимемо усну природну мову за допомогою семантичної мережі. Використаємо найпростіший метод. Всі можливі речення

розміститимо в предметних полях. В свою чергу всі речення кожного предметного поля (ПП) розділимо на категорії на основі смислів, що передаються. Кожному предметному полю відповідає порівняно невелике число категорій смислів (КС).

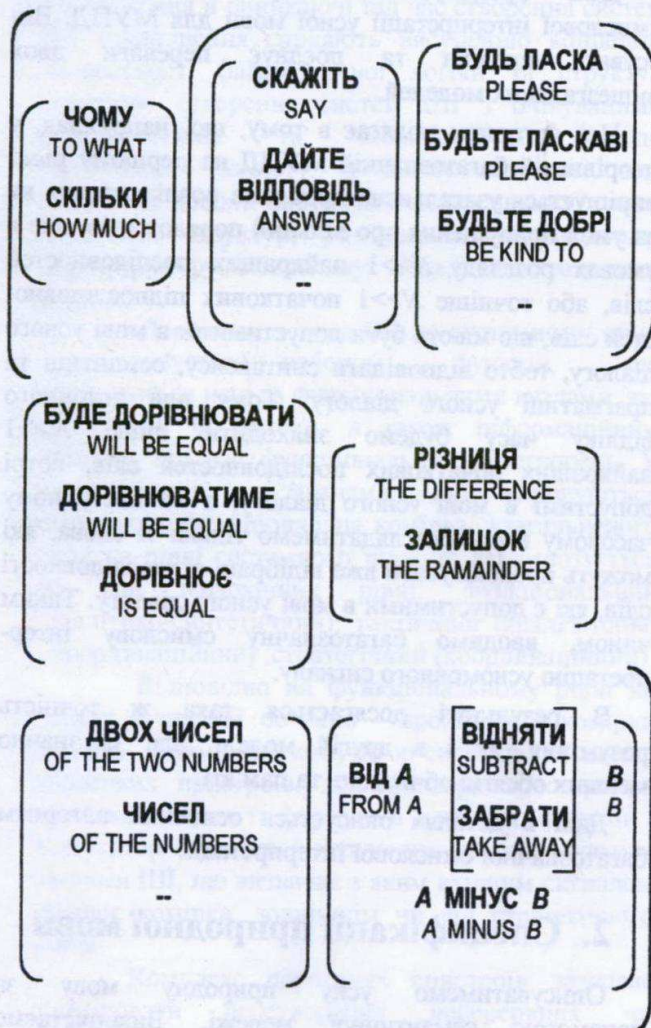
Наприклад, для довідкової служби аеропорту можна виділити такі категорії смислів: питання про приліт та відліт літаків; питання про наявність квитків; питання стосовно маршруту; питання про розташування служб аеропорту тощо.

Кожна категорія смислів (КС) складається з власної множини типів речень. Тип речення (ТР) — це конструкція, що економно задає множину речень, які отримуються з одного речення шляхом незалежних замінів та переставлянь як окремих слів, так і словосполучень. Базовим елементом ТР є підсловник. Підсловники в ТР іменуються згідно семантики предметного поля.

Кожна КС має порівняно невелику кількість ТР. Вочевидь, КС можуть при потребі поповнюватися новими ТР.

Всі ТР легко задаються списочними мовами, наприклад *LISP*.

Наводимо приклад ТР для питань, що стосуються різниці двох чисел для української мови:



Круглі дужки () містять підсловники, які можна

переставляти, тоді як квадратні [] містять неінверсні підсловники. Підсловники можна переставляти лише в межах "старших" дужок (). Як правило, ТР параметризуються. В цьому прикладі параметрами є операнди *A* і *B*. Символ "--" означає порожнє слово.

Неважко переконатися, що навіть якщо не брати до розгляду розмаїтість операндів *A* і *B*, наведений ТР задає загалом $6! \cdot (2 \cdot (2 \cdot (3 \cdot 4))) \cdot 3 \cdot 2 \cdot 3 \cdot 3 = 1\,866\,240$ різних допустимих в усній українській мові речень стосовно питання про різницю двох чисел. Серед них, зокрема, такі речення:

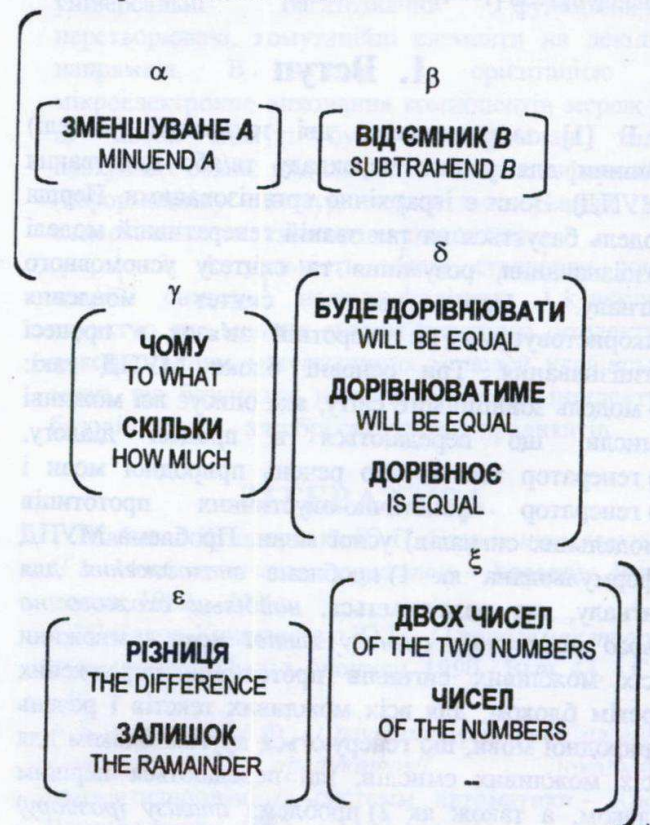
ЧОМУ СКАЖІТЬ ДОРІВНЮЄ РІЗНИЦЯ ЧИСЕЛ А МІНУС В,

ЧОМУ ДОРІВНЮВАТИМЕ РІЗНИЦЯ А МІНУС В СКАЖІТЬ БУДЬТЕ ЛАСКАВІ,

СКІЛЬКИ ВІД А ВІДНЯТИ В БУДЕ ДОРІВНЮВАТИ ЗАЛИШОК ДВОХ ЧИСЕЛ,

ВІД А ВІДНЯТИ В ЧОМУ ДОРІВНЮЄ ЗАЛИШОК ДАЙТЕ ВІДПОВІДЬ.

Нижче наведено інший приклад ТР для КС стосовно різниці двох чисел:



Для цього ТР підсловники іменуються так: α — зменшуване, β — від'ємник, γ — питальне слово, δ позначає дію, ϵ — операція і ξ — об'єкт дії.

КС і ТР використовуватимуться в процесі багатозначної смислової інтерпретації злитого мовлення. Тут слід підкреслити, що структури ТР зручні для генерування слів, що продовжують допустимі початкові підпоследовності слів.

3. Узагальнена проблема розпізнавання послідовності слів для злитого мовлення

Спершу розглянемо узагальнену проблему розпізнавання злитого мовлення, що складається зі слів вибраного словника [1, 2]. Потім цей результат узагальнимо для знаходження $N \gg 1$ смислів, що передаються злитим мовленням [3, 4].

3.1. Постановка задачі

Узагальнена проблема розпізнавання злитого мовлення полягає в тому, що виходячи з припущення про вільний порядок слів, знаходяться $N \gg 1$ найкращих різних послідовностей слів, впорядкованих за спаданням подібності до оброблюваного сигналу.

Нехай задані такі дані та знання:

А. Скінченна множина E елементарних прототипів усномовного сигналу (ЕПУС) $e(k^1) \in E$, де $k^1 \in K^1$ — ім'я ЕПУСу в алфавіті імен K^1 . Наприклад, маємо $|K^1| = |E| = 2^{10}$ елементів у E і K^1 . Отже, множина K^1 складає 1-й рівень модельних сигналів (мікрофонем), а пара (K^1, E) є кодовою книгою.

В. Скінченна множина K^2 — множина фонем-трифонів (ФТ) $k^2 \in K^2$ (ФТ формують 2-й рівень ієрархії модельних сигналів). ФТ є базовою фонемою, що розглядається в контексті впливу сусідніх фонем: першої, що передує, і другої, наступної. Для кожної природної мови фіксується біля 2000—3000 базових ФТ. Кожний ФТ k^2 з K^2 задається своєю транскрипцією в алфавіті K^1 :

$$k^2 = (k_1^1, k_2^1, \dots, k_s^1, \dots, k_{q(k^2)}^1),$$

де s позначає порядковий номер в транскрипції та $q(k^2)$ — тривалість транскрипції для k^2 .

С. Словник K^3 слів $k^3 \in K^3$, кожне з яких описане фонемною транскрипцією k^3 або фонемно-трифонною транскрипцією k^3 , заданою в алфавіті K^2 :

$$k^3 = (k_1^2, k_2^2, \dots, k_s^2, \dots, k_{q(k^3)}^2),$$

де $q(k^3)$ — довжина транскрипції слова.

Д. Розподіли $P(x/k^1)$ спостережуваних елементів x для всіх $k^1 \in K^1$, зокрема для всіх $k^1 \in K^1$:

$$P(x/k^1) = P(x/e(k^1)).$$

Дані та знання, розглянуті в А, В, С, Д, знаходяться в режимі навчання розпізнаванню [3].

При цьому формується так званий усномовний файл диктора.

Сигнал, який слід розпізнати, позначається послідовністю X_{0l} спостережуваних елементів x_i :

$$X_{0l} = (x_1, x_2, \dots, x_i, \dots, x_l),$$

де елементи x_i спостерігаються в рівновіддалених або майже рівновіддалених часових відліках i , а l — тривалість сигналу. Сегмент

$$X_{uv} = (x_{u+1}, x_{u+2}, \dots, x_v), \quad 0 \leq u < v \leq l$$

розглядається як реалізація сигналу фонем-трифона

$k^2 = (k_1^1, k_2^1, \dots, k_s^1, \dots, k_{q(k^2)}^1)$ з імовірністю

$$P(X_{uv}/k^2) = \max_{\{t_s\}} \prod_{s=1}^{q(k^2)} \prod_{i=t_{s-1}+1}^{t_s} P(x_i/k_s^1),$$

де $t_0 = u$, $t_{s-1} < t_s$, $t_{q(k^2)} = v$. Отже, ця імовірність обчислюється згідно з фонемно-трифонною транскрипцією k^2 та як згортка за границями мікрофоном $\{t_s\}$.

Стохастична автоматна породжувальна граматику (граф) для порівняння спостережуваного сегмента X_{uv} з усіма прототипами, що генеруються для ФТ k^2 , зображена на рис. 1а. Цей граф має $q(k^2)$ станів. Кожному стану s приписується мікрофонема k_s^1 або $k^1(s)$. Переходи здійснюються згідно стрілок за 0 або 1 часових кроків.

Зображуючи граф фонем-трифона пунктирним прямокутником і об'єднуючи графи фонем-трифонів в лінійну послідовність згідно транскрипції слова k^3 , отримуємо стохастичну автоматну породжувальну граматику для генерування прототипів (модельних сигналів) слів і їх порівняння з розпізнаваним сигналом. Граф слова зображено як "пелюстку" "квітки" на рис. 1б. Граф-квітка на рис. 1б представляє стохастичну автоматну породжувальну граматику для складання модельних сигналів злитого мовлення за умови вільного порядку слів. Кількість пелюсток збігається з кількістю слів у словнику.

Імовірність сегмента X_{uv} за умови слова k^3 обчислюється згідно з транскрипцією слова $k^3 = (k_1^2, k_2^2, \dots, k_s^2, \dots, k_{q(k^3)}^2)$ як згортка за границями фонем-трифонів $\{w_s\}$:

$$P(X_{uv}/k^3) = \max_{\{w_s\}} \prod_{s=1}^{q(k^3)} P(X_{w_{s-1}w_s}/k_s^2),$$

де $w_0 = u$, $w_{s-1} < w_s$, $w_{q(k^3)} = v$.

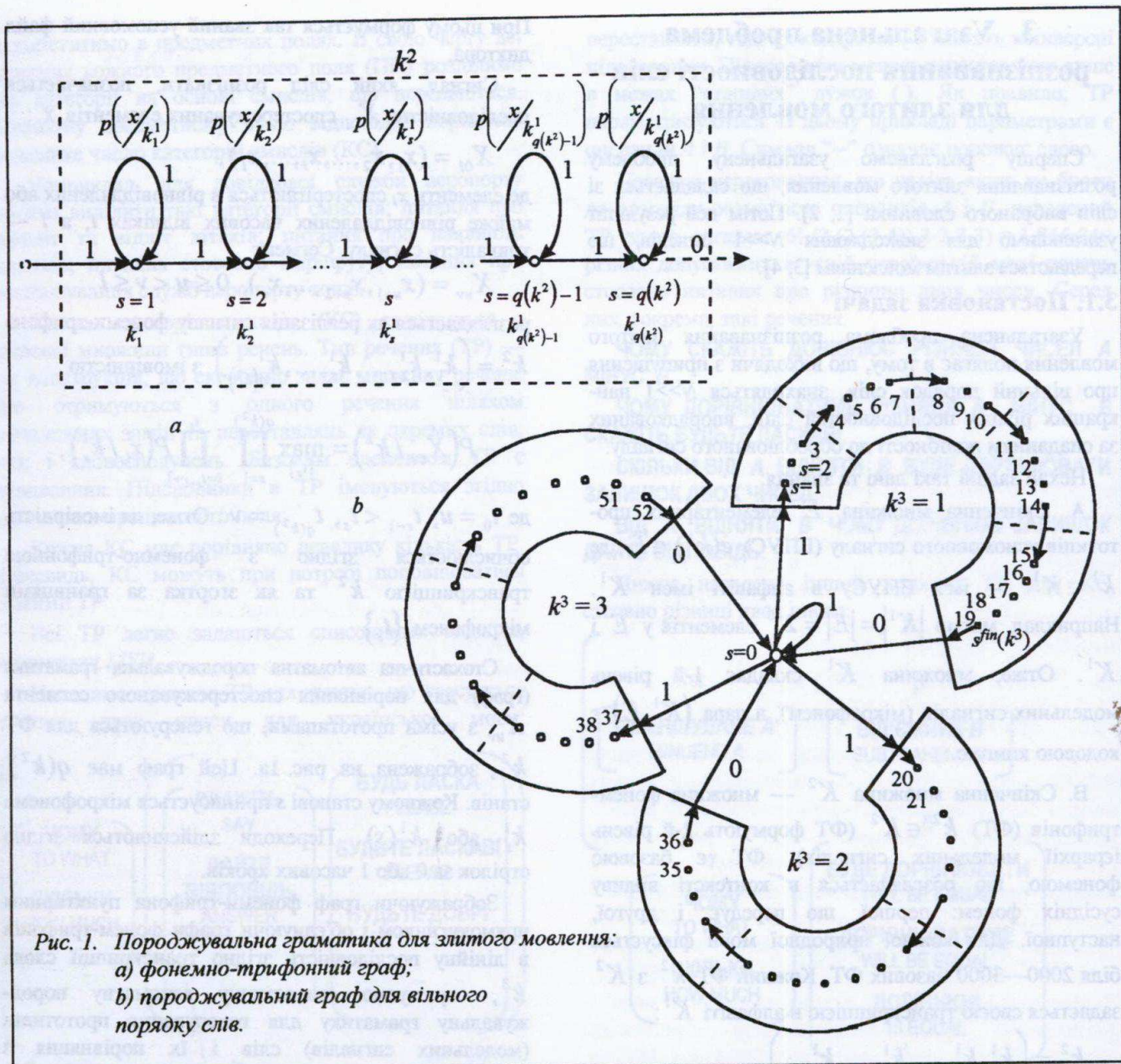


Рис. 1. Породжувальна граматика для злитого мовлення:
 а) фонемно-трифонний граф;
 б) породжувальний граф для вільного порядку слів.

Відповідно, найкраща відповідь розпізнавання у вигляді послідовності слів за умови їх вільного порядку і невідомої кількості слів Q у послідовності визначається максимізацією наступного виразу правдоподібності (1):

$$P(X_{0i}/k_1^3, \dots, k_s^3, \dots, k_Q^3) = \max_{\{m_s\}} \prod_{s=1}^Q P(X_{m_{s-1}m_s}/k_s^3)$$

де $m_0 = 0$, $m_{s-1} < m_s$, $m_{q(k^3)} = l$ — границі слів в усномовному сигналі X_{0i} .

3.2. Узагальнений алгоритм розпізнавання усної мови

Щоб сформулювати узагальнений алгоритм розпізнавання і таким чином знайти $N \gg 1$ найкращих послідовностей слів, введемо наскрізну нумерацію станів для графа злитого мовлення (див. рис. 1б). Розрізнятимемо головний стан $s=0$ для

паузи, а також опустимо індекс 3 в записі k^3 . Отже, лінійки станів $s=1:19$, $s=20:36$ і $s=37:52$ представляють слова $k=1$, $k=2$ і $k=3$ відповідно.

Окремо виділимо вхідні стани слів $s^{in}(k)$ і вихідні $s^{fin}(k)$. Це $s=1, 20, 37$ and $s=19, 36, 52$ відповідно.

Нехай $\Omega_i(s)$ позначає множину модельних сигналів злитого мовлення довжини i , що генеруються графом злитого мовлення в результаті переходів зі стану $s=0$ до стану s за i кроків часу. Позначимо через $(F_i^r(s), \mathcal{K}_i^r(s))$, $r=1:N$ N -ку найкращих імовірностей $F_i^r(s)$, $r=1:N$ (1), що досягаються на множині $\Omega_i(s)$, але для початкового сегмента сигналу $X_{0i} = (x_1, x_2, \dots, x_i)$, а через $\mathcal{K}_i^r(s)$, $r=1:N$ позначимо відповідну оптимальну

підпоследовність слів. В N -ці $(F_i^r(s), \mathcal{K}_i^r(s))$, $r = 1:N$ всі $\mathcal{K}_i^r(s)$ є різними. Нехай пари $(F_i^r(s), \mathcal{K}_i^r(s))$ в N -ці $(F_i^r(s), \mathcal{K}_i^r(s))$, $r = 1:N$ розташовані в порядку спадання величини $F_i^r(s)$:

$$F_i^1(s) \geq F_i^2(s) \geq \dots \geq F_i^r(s) \geq \dots \geq F_i^N(s).$$

Нехай N -ка $(F_v^r(s), \mathcal{K}_v^r(s))$, $r = 1:N$ вже обчислена для всіх станів s і для всіх відліків часу $v < i$, котрі передують i .

Тоді після появи наступного елемента x_i в момент часу i для всіх станів s обчислюється нова N -ка $(F_i^r(s), \mathcal{K}_i^r(s))$, $r = 1:N$:

а) по-перше, для всіх внутрішніх станів слів s (окрім вхідних станів слів $s^{in}(k)$, це $s=1, 20, 37$ на рис. 1б, і головного стану $s=0$):

спочатку обчислюються всі $2N$ можливі добутки $F_{i-1}^w(s-1)P(x_i/k^1(s))$, $F_{i-1}^t(s)P(x_i/k^1(s))$, $w, t = 1:N$ і виписуються відповідні $2N$ послідовностей слів $\mathcal{K}_{i-1}^w(s-1)$, $\mathcal{K}_{i-1}^t(s)$, $w, t = 1:N$, і тоді формується нова N -ка $(F_i^r(s), \mathcal{K}_i^r(s))$, $r = 1:N$ шляхом вибору N найкращих добутків $F_i^r(s)$, що впорядковуються, з відповідними різними $\mathcal{K}_i^r(s)$ у цій новій N -ці;

б) по-друге, для всіх вхідних станів слів $s^{in}(k)$, $k \in K$, це $s=1, 20, 37$ на рис. 1б:

спочатку обчислюються всі $2N$ можливі добутки $F_{i-1}^w(0)P(x_i/k^1(s))$, $F_{i-1}^t(s^{in}(k))P(x_i/k^1(s))$, $w, t = 1:N$ і виписуються відповідні $2N$ послідовностей слів $\mathcal{K}_{i-1}^w(0)$, $\mathcal{K}_{i-1}^t(s^{in}(k))$, $w, t = 1:N$, і тоді нова N -ка $(F_i^r(s), \mathcal{K}_i^r(s))$, $r = 1:N$ формується шляхом вибору N найкращих впорядкованих добутків $F_i^r(s)$ з відповідно різними $\mathcal{K}_i^r(s)$ в цій новій N -ці;

с) по-третє, для головного стану $s=0$:

спершу виписуються та обчислюються всі можливі $(|K|+1)N$ значення $F_i^w(s^{fin}(k))$, $w = 1:N$, $k = 1:|K|$ та $F_{i-1}^u(0)P(x_i/k^1(0))$, $u = 1:N$ ($s^{fin}(k)$ є вихідними станами слів, це $s=19, 36, 52$ на рис. 1б) і формуються $(|K|+1)N$ нових послідовностей слів $\mathcal{K}_i^w(s^{fin}(k)) \oplus k$, $w = 1:N$, $k = 1:|K|$ і виписуються $\mathcal{K}_{i-1}^u(0)$,

$u = 1:N$, відповідно (тут позначка \oplus означає додавання нового слова до послідовності слів), і тоді нова N -ка $(F_i^r(0), \mathcal{K}_i^r(0))$, $r = 1:N$ знаходиться шляхом вибору N найкращих і впорядкованих $F_i^r(0)$ з відповідно різними $\mathcal{K}_i^r(0)$ в новій N -ці.

Узагальнена відповідь розпізнавання визначається N -кою, обчисленою для головного стану $s=0$ в момент часу $i=l$: $\mathcal{K}_i^r(0)$, $r = 1:N$.

На початку алгоритму покладаємо: $F_0^1(0) = 1$, $F_0^r(0) = 0$ для $r = 2:N$, $F_0^r(s) = 0$ для всіх $r = 1:N$ і всіх $s \neq 0$, $\mathcal{K}_0^r(s) = \emptyset$ для всіх $r = 1:N$ і всіх s .

Неважко помітити, що поки розглядаються внутрішні стани слова, нові слова лише "розгортаються", а не додаються до підпоследовностей слів $\mathcal{K}_i^r(s)$, тоді як тільки при переході з вихідного стану слова $s^{fin}(k)$ до головного стану $s=0$ трапляється, що нові слова додаються до вже накопичених підпоследовностей слів $\mathcal{K}_i^r(s)$.

4. Алгоритм багатозначної смислової інтерпретації

Алгоритм багатозначного розпізнавання злитого мовлення подано в такій формі, що робить можливим поширення цього алгоритму на смислову інтерпретацію злитого мовлення.

Використовуватимемо поняття типів речень (ТР) і категорій смислів (КС), введених в розділі 2. Зокрема, структура ТР дає просте правило розпізнавання, чи може якесь слово продовжити вже накопичену послідовність слів, або чи є допустимим певна послідовність слів для даного ТР.

Нехай $\Lambda(\mathcal{K}_i^r(s))$ — підсловник слів, що можуть продовжити початкову послідовність слів $\mathcal{K}_i^r(s)$. Коли $\mathcal{K}_i^r(s)$ заповнена і не може бути продовжена, тоді $\Lambda(\mathcal{K}_i^r(s))$ дорівнює всьому словнику K , оскільки можливе започаткування нових послідовностей слів.

Аналізуючи підпоследовність слів $\mathcal{K}_i^r(s)$ за всіма структурами ТР і КС, неважко знайти послідовність смислів $M(\mathcal{K}_i^r(s))$, що передається сигналом злитого мовлення X_{0i} .

Тепер, посилаючись на розділ 3, стисло подамо алгоритм багатозначної смислової інтерпретації злитого мовлення.

Нехай $(F_i^r(s), \mathcal{K}_i^r(s), M_i^r(s)), r = 1: N$ є N -кою найкращих імовірностей $F_i^r(s), r = 1: N$ (1), котрі досягаються на множині прототипів злитого мовлення $\Omega_i(s)$ для початкового сигналу $X_{0i} = (x_1, x_2, \dots, x_i)$, і $\mathcal{K}_i^r(s), r = 1: N$ — відповідні підпоследовності слів, що є різними, і $M_i^r(s) = M(\mathcal{K}_i^r(s)), r = 1: N$ — відповідні оптимальні результати багатозначної смислової інтерпретації $(F_i^r(s), \mathcal{K}_i^r(s), M_i^r(s)), r = 1: N$ в N -ці, впорядкованій за спаданням $F_i^r(s)$.

Нехай N -ки $(F_v^r(s), \mathcal{K}_v^r(s), M_v^r(s)), r = 1: N$ вже обчислені для всіх станів s і для всіх відліків часу $v < i$, що передують i .

Тоді після появи наступного спостережуваного елементу x_i в момент часу i для всіх станів s обчислюється нова N -ка $(F_v^r(s), \mathcal{K}_v^r(s), M_v^r(s)), r = 1: N$:

а) по-перше, для внутрішніх станів слів s спочатку обчислюються всі можливі $2N$ добутки $F_{i-1}^w(s-1)P(x_i / k^1(s)), F_{i-1}^t(s)P(x_i / k^1(s)), w, t = 1: N$ і виписуються відповідні $2N$ послідовності $(\mathcal{K}_{i-1}^w(s-1), M_{i-1}^w(s-1)), (\mathcal{K}_{i-1}^t(s), M_{i-1}^t(s)), w, t = 1: N$ слів і смислів, і тоді нова N -ка $(F_i^r(s), \mathcal{K}_i^r(s), M_i^r(s)), r = 1: N$ формується шляхом вибору N найкращих добутків $F_i^r(s)$, що впорядковуються, з відповідно різними $\mathcal{K}_i^r(s)$ в цій новій N -ці;

б) по-друге, для всіх вхідних станів слів $s^{in}(k), k \in K$ спершу обчислюються всі можливі $2N$ добутки $F_{i-1}^w(0)P(x_i / k^1(s)), F_{i-1}^t(s^{in}(k))P(x_i / k^1(s)), w, t = 1: N$, і виписуються відповідні $2N$ послідовності слів і смислів $(\mathcal{K}_{i-1}^w(0), M_{i-1}^w(0)), (\mathcal{K}_{i-1}^t(s^{in}(k)), M_{i-1}^t(s^{in}(k))), w, t = 1: N$, і тоді нова N -ка $(F_i^r(s), \mathcal{K}_i^r(s), M_i^r(s)), r = 1: N$ формується шляхом вибору N найкращих і впорядкованих добутків $F_i^r(s)$ з відповідними різними $\mathcal{K}_i^r(s)$ в цій новій N -ці;

в) по-третє, для головного стану $s=0$ спершу виписуються та обчислюються всі можливі $(|K|+1)N$ значень $F_i^w(s^{in}(k)), w = 1: N, k = 1: |K|$ і $F_{i-1}^u(0)P(x_i / k^1(0)), u = 1: N$, і формуються та беруться до уваги всі можливі нові підпо-

слідовності слів і результатів смислової інтерпретації $(\mathcal{K}_i^w(s^{in}(k)) \oplus k, M(\mathcal{K}_i^w(s^{in}(k)) \oplus k)), w = 1: N, k \in \Lambda(\mathcal{K}_i^w(s^{in}(k)))$ і $(\mathcal{K}_{i-1}^u(0), M_{i-1}^u(0)), u = 1: N$ відповідно, і тоді знаходиться нова N -ка $(F_i^r(0), \mathcal{K}_i^r(0), M_i^r(0)), r = 1: N$ шляхом вибору N найкращих і впорядкованих $F_i^r(0)$ з відповідними різними $\mathcal{K}_i^r(0)$ в цій новій N -ці.

Відповідь багатозначної смислової інтерпретації визначається N -кою, обчисленою для головного стану $s=0$ в момент часу $i=l$: $(\mathcal{K}_i^r(0), M_i^r(0)), r = 1: N$.

На початку алгоритму багатозначної смислової інтерпретації покладаємо: $F_0^1(0) = 1, F_0^r(0) = 0$ for $r = 2: N, F_0^r(s) = 0$ для всіх $r = 1: N$ і всіх $s \neq 0, \mathcal{K}_0^r(s) = \emptyset$ та $M_0^r(s) = \emptyset$ для всіх $r = 1: N$ і всіх s .

5. Прикінцеві положення

Алгоритм багатозначної смислової інтерпретації злитого мовлення оперує лише з допустимими в діалозі початковими підпоследовностями слів, і навіть більше: щоб уникнути факту локального рішення, багатозначний розв'язок $N \gg 1$ вводиться в кожний момент часу.

Очевидно, що більш раціонально змінювати кількість багатозначних рішень N , починаючи з великих чисел, зменшуючи їх в процесі накопичення інформації при обробленні усномовного сигналу.

Алгоритм багатозначної смислової інтерпретації злитого мовлення не гарантує глобального розв'язку проблеми розуміння, але за певного вибору кількості смислів N в реальному комп'ютерному середовищі знаходиться прийнятний прагматичний результат.

ЛІТЕРАТУРА

1. Т.К. Vintsiuk. *Speech Recognition and Understanding*. — Kibernetika, 1982, No. 5, pp 101-111.
2. Т.К. Винцюк. *Обобщенная задача распознавания слитной речи*. — Труды VIII Всесоюзного семинара "Автоматическое распознавание слуховых образов 1982", Киев, 1982, с. 345-348.
3. Т.К. Винцюк. *Анализ, распознавание и смысловая интерпретация речевых сигналов*. — Киев: Наукова думка, 1987, 264 с.
4. Taras K. Vintsiuk. *Two Approaches to Create a Dictation/Translation Machine*. Proceedings of the Second International Workshop "Speech and Computer", Cluj-Napoca, 1997, pp 1-6.