## Зміст

Зміст 2

Передумови створення морфологічних процесорів 3

Загальна форма морфологічного аналізу текстів 4

Деякі обмеження 4

Термінологія 4

Основні моменти 7

Загальний вигляд алгоритму морфологічного аналізу 10

Внутрішня будова алгоритму 10

Частина А – “Загальні правила” 10

Частина Б – “Список суфіксів” 12

Частина В – “Список інформацій до суфіксів” 13

Частина Г – “Нестандартний запис” 14

Частина Д – “Омонімія словоформ” 14

Висновки 15

Література 16

## Передумови створення морфологічних процесорів

Припустимо, що весь процес аналізу під час автоматичного перекладу поділяється на два основних етапи: морфологічний аналіз та синтаксичний аналіз.

Під морфологічним аналізом розуміють повну обробку взятих окремо (без будь-якого зв’язку з контекстом) словоформ. В результаті такої обробки кожна словоформа замінюється так званою інформацією – ланцюжком символів, що визначають усі ті властивості даної словоформи, що необхідні для подальшого аналізу (і подальшого перекладу). Інформації до словоформ використовуються на другому етапі аналізу – під час синтаксичного аналізу, в результаті якого встановлюються зв`язки між усіма словоформами тексту та між частинами складних речень.

Морфологічний аналіз – основний інструмент морфологічних процесорів. Створення повнофункціонального морфологічного процесору неможливе без детального розуміння принципів морфологічного аналізу.

За роки існування автоматичного перекладу було створено алгоритми морфологічного аналізу для багатьох мов. Природно, що ці алгоритми мають багато спільного. Однак мають місце також і значні розбіжності, що зумовлені властивостями будови відповідних мов, а також різними підходами.

Виникає природне бажання узагальнити досвід, що накопичено в цій галузі. Зокрема, бажано відповісти на таке питання: чи не можна виділити в алгоритмах морфологічного аналізу стандартні частини, що повторюються в багатьох алгоритмах? Чи не можна на базі стандартних частин, що виділено, розробити для алгоритму морфологічного аналізу таку спільну форму, щоб вона підходила (без суттєвих змін) для багатьох мов? Чи не можна, зрештою, встановити зв`язок між властивостями конкретних мов та деталями алгоритмів морфологічного аналізу таким чином, щоб загальну форму алгоритму можна було пристосувати до якоїсь конкретної мови відповідно до правил, що вказано заздалегідь (ці правила можуть бути приблизно такими: “За певної властивості мови із загальної форми треба(не треба) взяти певний елемент і т. ін.”)?

Нижче робиться спроба дати позитивну відповідь на ці питання.

## Загальна форма морфологічного аналізу текстів

### Деякі обмеження

Загальна форма морфологічного аналізу, про яку іде мова, не є абсолютно загальною. Її “загальність” обмежено певним колом мов та певним типом алгоритмів.

Обмеження стосовно мов полягає в тому, що загальна форма морфологічного аналізу призначена лише для суфіксально-афіксальних мов (індоєвропейських, угро-фінських, тюркських і т. ін.). Виключаються мови з розвиненою внутрішньою флексією (семітські) та інфіксацією (на зразок низки індонезійських).

Окрім того, для мов, у яких словоформи взагалі не поділяються на морфеми (так звані “ізолюючі” – на зразок в`єтнамської), морфологічний аналіз повністю зводиться до пошуку в словнику, а решта операцій, що передбачені у загальній схемі алгоритму морфологічного аналізу, для “ізолюючих” мов не потрібні.

Обмеження стосовно типу алгоритмів полягає в тому, що загальна форма морфологічного аналізу охоплює лише такі алгоритми, в основі яких полягає розбиття словоформ на морфеми від початку до кінця (для мов з латинською та кириличною графікою – зліва направо). Можливий інший підхід, за якого розбиття словоформ починається з кінця. Найкращим розв`язком, очевидно, було б узагальнення та гнучке поєднання обох підходів.

### Термінологія

Перш ніж перейти до опису загальної схеми, зупинимося на деяких термінах, якими нам доведеться користуватися надалі.

Як вже говорилося, морфологічний аналіз – це обробка взятих окремо словоформ. Морфологічний аналіз полягає в такому: словоформа розбивається на морфеми (за умови, якщо таке розбиття можливе) шляхом пошуку морфем у спеціальних списках, де кожній морфемі відповідає певна інформація; потім із інформацій до морфем будується загальна інформація до словоформи.

В даному описі морфема не визначається: це об`єкт окремого дослідження. Морфеми подаються списком. В алгоритмі морфологічного аналізу, що описується, як морфеми взято практично ті самі елементи, що їх вважають морфемами в лінгвістиці. Однак слід мати на увазі основні відмінності двох типів:

1. Оскільки ми маємо справу з письмовим текстом, а більшість лінгвістичних праць – з усною мовою, то поділ на морфеми, що здійснюється для алгоритму морфологічного аналізу, інколи не співпадає з тим, що прийнятий в лінгвістиці.

Розбіжність між написанням та вимовою (та, зокрема, відсутність наголосу у письмовому тексті) відбивається також на зовнішньому (графічному чи звуковому) боці морфем. Морфеми, що відрізняються графічно, можуть мати однакове звучання: *-ом* (ор. відм. однини) та *–ам* (дав. відм. множини) в таких словоформах, як *лаком – лакам* або *муром – мурам*. З іншого боку, морфеми, що мають різне звучання, можуть співпадати графічно: *-и* (род. відм. однини) та *–и* (наз.-знах. відм. множини) в таких словоформах, як *казки – казки*.

Алгоритм працює лише з графічними формами морфем.

1. Оскільки в багатьох лінгвістичних працях під час поділу словоформ на морфеми не завжди послідовно розрізняють синхронічний та діахронічний аспекти, прийнятий поділ, що виправданий в низці відношень, може виявитися незручним з точки зору автоматичного аналізу та перекладу. В таких випадках доводиться відступати від прийнятого поділу. Наприклад, словоформа *відношення* зазвичай поділяється на *від-нош-енн-я*, а для алгоритму, що описується – на *відношенн-я.*

 Ми не будемо спеціально викладати тут тих міркувань, за якими приймається певний поділ тієї чи іншої конкретної словоформи. Запропонований алгоритм морфологічного аналізу (у загальній формі) абсолютно не залежить від того, які конкретні елементи обрано як морфеми. Якщо прийняти інший поділ на морфеми, це змінить лише склад списків морфем, але не сам алгоритм. А тому тимчасово можна вважати, що, не уміючи точно визначати морфеми, ми обмежуємося їх емпіричним підбором (базуючись на інтуїції та вимогах перекладу) та подаємо їх простим перерахуванням.

 Серед морфем розрізнятимемо основи та афікси.

 Терміни “основа” та “афікс” вживаються тут не в своєму традиційному значенні. Прийняте у лінгвістиці розрізняння коренів, основ, кореневих основ і т. ін. для наших цілей є несуттєвим, і все це позначається спільним терміном “основа”. Визначення цього терміну ми не даємо; ми подаємо основи списком. Однак, можна вказати такі властивості основ:

1. Будь-яка словоформа містить принаймні одну основу (може бути й більше).
2. Основи складають переважну більшість морфем і є численними. У загальнолітературній мові кількість основ сягає сотень тисяч; у текстах із спеціальних галузей знання вона зменшується до кількох тисяч (або десятків тисяч).
3. Одна й та сама основа зустрічається в порівняно невеликій кількості різних словоформ – від однієї до кількох десятків.
4. Від основи до загальної інформації до словоформи надходять значення більшості ознак, що складають цю інформацію.

 Морфеми, що не є основами, називаються афіксами (тобто не проводиться традиційної різниці між афіксами та флексіями). Афікси, що зустрічаються перед основами, називаються префіксами; афікси, що зустрічаються позаду основ, називаються суфіксами. Таким чином, суфіксами вважаються не тільки морфеми словотворення, але також і так звані “тематичні голосні”, “з`єднувальні звуки” (а точніше – “літери”), флексії і т.ін.

 Властивості афіксів певним чином протилежні властивостям основ:

1. Афікси зустрічаються не у всіх словоформах.
2. Афікси завжди складають меншість морфем, їх загальна кількість невелика: від кількох десятків до кількох сотень (залежно від мови).
3. Один і той самий афікс зазвичай зустрічається в порівняно великій кількості різних словоформ – від кількох десятків до кількох тисяч і навіть десятків тисяч, але ніколи не менше, ніж у двох словоформах (основа може зустрічатися і в одній словоформі, співпадаючи з нею).
4. Від афікса до загальної інформації до словоформи надходять значення меншості ознак, що складають цю інформацію.

 В даному описі афікси також подаються списком.

 В процесі розробки алгоритмів, а саме – під час складання словника, зручно задавати лише афікси (оскільки у мовах, які ми розглядаємо, їх кількість відносно невелика) та, виходячи з прийнятого списку афіксів, виділяти основи (з поданих у тексті словоформ) за допомогою певних правил.

 Частина словоформи, що залишається після відкидання будь-яких морфем від початку словоформи, називається залишком. Під час аналізу залишок весь час зменшується по мірі послідовного відтинання морфем – до повного зникнення, що означає кінець аналізу для даної словоформи. Наприклад, під час аналізу словоформи *перекомпілювати* після відтинання префікса *пере-* залишком буде *–компілювати*, після відтинання основи *компіл-* залишається *–ювати*, після відтинання суфікса *–юва–*  залишок буде *–ти,* і зрештою, після відтинання суфікса *–ти* залишок зникає і аналіз закінчується.

 Дві морфеми, що містять різну інформацію, можуть графічно співпадати (повністю або частково) одна з одною. Ми ж роздивимося лише такі випадки співпадання, коли принаймні одна з морфем повністю входить до складу іншої. Решта випадків співпадання не впливають на правильність морфологічного аналізу й тому нас не цікавлять.

 Як приклад повної вкладеності одної морфеми в іншу можна навести суфікс *–ами* (ор. відм. мн.), що повністю вкладається в основу *намист-* слова *намисто*, що призводить до неправильного розкладу *н-ами-ст-о*

 Особливо важливі два приватних випадки співпадання морфем.

 а) Дві морфеми повністю співпадають. Такі морфеми називаються *омонімічними.*

б) Одна з морфем входить до складу іншої, і при цьому існує деяка одиниця, що починається з меншої морфеми, але дозволяє вкладення і більшої (тої, що включає) морфеми. Тоді ми кажемо, що ці морфеми *спряжені* одна з одною.

Вкладення одних морфем у інші, і, зокрема омонімія та спряженість морфем створюють певні труднощі морфологічного аналізу.

### Основні моменти

У процесі морфологічного аналізу легко виділяються такі основні моменти, що є спільними для різних мов та алгоритмів:

1. Пошук словоформи (або залишку) в певних списках морфем (у таблицях). Під пошуком мається на увазі послідовне порівняння текстової одиниці (словоформи або залишку) з елементами списку (з основами, префіксами або суфіксами) доти, доки елемент списку політерно не “вкладеться” зліва направо до текстової одиниці, що обробляється.
2. Вибір потрібного елемента з кількох можливих. Тут мається на увазі випадок, коли до текстової одиниці одночасно можуть вкладатися різні морфеми (омонімічні або спряжені). Необхідний вибір робиться на основі спеціальних поміток, що приписані морфемам та відображають їхню сполучність. Так, за допомогою поміток при основі обирається правильний суфікс (із декількох, що вкладаються до цього закінчення); при розтинанні залишку на суфікси може бути виправлено помилку, що її припустилися раніше (під час відтинання основи) і т.д.
3. Вибір інформацій до окремих морфем та об`єднання цих інформацій у загальну інформацію до словоформи, що аналізується.
4. Виявлення індивідуальних особливостей морфем (якщо такі особливості є) та врахування впливу цих особливостей на загальну інформацію до словоформи.
5. Вибір подальших дій після того, як завершено обробку наступної морфеми. Послідовність дій зазвичай визначається властивостями щойно обробленої морфеми.

 Конкретний зміст цих п`яти основних моментів морфологічного аналізу змінюється в залежності від алгоритму, але самі ці моменти залишаються в принципі незмінними. Не обов`язково всі п`ять моментів мають місце: другий та четвертий можуть взагалі бути відсутні (наприклад, у мові, де немає морфем, що співпадають і де морфеми не мають індивідуальних особливостей). Важливо лише те, що для опису морфологічного аналізу вищенаведених моментів досить.

 Найбільш складним із вказаних моментів морфологічного аналізу є другий – обрання потрібної морфеми у випадку омонімії, повної вкладеності та спряженості морфем. Якби в мовах не було всіх цих “небезпечних ” спів падань морфем, то морфологічний аналіз здійснювався б кількома стандартними, надзвичайно простими операціями. Взагалі, головна скланість автоматичного перекладу в цілому – це зовнішнє (в даному випадку, графічне) співпадання мовних елементів, що по-різному перекладаються й тому містять різну інформацію. Таке спів падання призводить до задачі розрізняння та вибору, без якої аналіз під час автоматичного перекладу зводився б до елементарних перетворень.

 Природно. що проблема омонімії (в найширшому розумінні терміна), тобто задача розрізняння та вибору, посідає важливе місце в морфологічному аналізі. Морфологічний аналіз має забезпечувати правильний розбір таких випадків співпадання, які принципово (людиною) можуть бути проаналізовані під час розгляду ізольованих словоформ, без звертання до контексту. Повинно правильно розкладатися співпадання окремих морфем за умови, що не співпадають цілі словоформи. Не можуть і не повинні однозначного розкладатися випадки, коли має місце повне співпадання словоформ. Вважається за достатнє, коли всі такі випадки буде якимось чином виділено для подальшого розкладу.

 Основні моменти морфологічного аналізу, що наведені вище, мають такий загальний характер, що самі по собі вони є тривіальними. Однак вони дозволяють мислити весь процес морфологічного аналізу чітко та компактно, хоча й у загальному вигляді. Для практичної розробки морфологічного аналізу ці основні моменти необхідно конкретизувати, точно описавши окремі алгоритми, за допомогою яких вони будуть здійснюватися і які всі разом складуть алгоритм морфологічного аналізу. Необхідно побудувати якийсь загальний алгоритм морфологічного аналізу, або розробити таку загальну форму для алгоритмів морфологічного аналізу, яка виявиться достатньою (хоча деколи й надто розширеною) для багатьох мов, що вже існують, а також теоретично можуть існувати.

 Нижче описано один із можливих варіантів такого загального алгоритму морфологічного аналізу.

 Однак слід мати на увазі, що загальну схему морфологічного аналізу викладено тут не повністю. Річ у тім, що, хоча пошук словоформ у словнику основ принципово не відрізняється від пошуку залишків у таблиці суфіксів, пошук у словнику основ являє собою певною мірою самостійну задачу. Це пов`язано з численністю основ. Великий обсяг словника основ (порівняно до таблиць суфіксів) зумовлює специфічні умови пошуку в ньому, вилучення інформації тощо. Тому вважається за доцільне умовно поділити морфологічний аналіз на два алгоритми: “Пошук у словнику основ” та “Власне морфологічний аналіз”. Тут взагалі не розглядається перший алгоритм. Не розглядається також відтинання префіксів, оскільки воно тісно пов`язане з пошуком у словнику основ і є

 розділом цієї частини аналізу. Мова йде про морфологічний аналіз з того моменту, коли від словоформи вже відділені префікси та одна основа. Маємо на увазі морфологічний аналіз відповідного залишку і вважаємо, що ми маємо словарну інформацію про основу, що відітнули (а також про префікси, якщо вони були).

###

### Загальний вигляд алгоритму морфологічного аналізу

 В загальному вигляді алгоритм морфологічного аналізу (без пошуку в словнику основ) складається з п`яти частин.

1. Частина А – “Загальні правила” – є основною частиною алгоритму (власне алгоритмом) і являє собою набір правил, що визначають послідовність операцій та взаємодію решти частин.
2. Частина Б – “Список суфіксів” – містить перелік суфіксів даної мови разом із деякими вказівками, що необхідні для їх правильної обробки (“допоміжними інформаціями”).
3. Частина В – “Список інформацій до суфіксів” (інакше – “Стандартний запис”), містить перелік “основних інформацій” до тих суфіксів, що містяться у частині Б (під “основною інформацію до суфіксу” мається на увазі та інформація, що її має бути перенесено до інформації до словоформи).
4. Частина Г – “Нестандартний запис” – є набором вказівок про індивідуальні особливості суфіксів співвідносно до особливостей основ а також про те, як ці особливості мають бути враховані.
5. Частина Д – це таблиця, де перераховано випадки, коли повністю співпадають деякі форми від різних основ (при тому, що інші форми від цих слів – різні).

 П`яти частин, що наведені вище, достатньо для представлення алгоритму морфологічного аналізу цілої низки мов. Частини А, Б, та В необхідні; частини Г та Д можуть бути відсутні.

 Частини алгоритму можуть поділятися на розділи.

 Перейдемо до опису внутрішньої будови частин алгоритму.

### Внутрішня будова алгоритму

# Частина А – “Загальні правила”

 Частина А являє собою певну послідовність правил, що викладені в термінах так званих алгоритмічних операторів. Оператор – це стандартна частина алгоритму, що повторюється, для якої можна створити стандартну програму реалізації.

 Частина А складається з десяти розділів; серед них розділи І. ІІІ, V та VI є обов`язковими для будь-якого алгоритму, а решта розділів може бути відсутня.

Кожний розділ має виконувати певну змістовну задачу, і його робота завжди призводить до певного результату. Для всіх розділів визначено, які результати (умовно) вважати позитивними, а які – негативними. Для кожного розділу також вказується , куди слід переходити у випадку того чи іншого результату.

Отже, частина А працює таким чином:

1. Під час морфологічного аналізу окремих словоформ відрізняють обробку “нормальних” випадків (розділи II-VI) та “особливих” випадків (розділи VII-X). “Нормальні” випадки – це словоформи, для яких було знайдено правильну основу та залишки яких (якщо залишки є) складаються лише із суфіксів, при цьому в цих суфіксах не мають місце графічні чергування. До “нормальних” випадків належать більшість словоформ тексту в таких мовах, як українська, російська, англійська, французька, угорська та ін. “Особливі” випадки – це або словоформи, для яких знайдено помилкову основу (внаслідок омонімії або спряженості основ), або словоформи із залишками, що містять, окрім суфіксів, ще й основу (“складні слова”), або словоформи, що мають графічні чергування в суфіксах. Словоформи, що відносяться до “особливих” випадків, зазвичай становлять меншість словоформ тексту.
2. Перед початком морфологічного аналізу будь-яка словоформа вважається “нормальною” і робиться спроба виконати “нормальну” обробку цієї словоформи. В більшості випадків це вдається зробити. Якщо ж під час “нормальної” обробки словоформа виявляється “ненормальною” (тобто один з розділів дає негативний результат), її передають на “особливу” обробку. Після “особливої” обробки (помилково знайдену словоформу замінюють іншою і т. д., див. нижче п. 5) словоформа повертається на “нормальну” обробку. Таким чином, “нормальна” обробка виконується над усіма словоформами, для яких потрібен морфологічний аналіз, а “особлива” – лише над деякими.
3. Обробка словоформ починається з вибору послідовності дій (розділ І): для деяких словоформ морфологічний аналіз взагалі не потрібен (так звані “незмінні слова”); для інших одразу з`ясовується необхідність “особливої” обробки, решта словоформ проходять “нормальну” обробку (при цьому під час аналізу вони можуть виявитися “особливими”).
4. “Нормальна” обробка складається з таких п`яти етапів:

 а) вибір таблиці суфіксів, у якій буде вестися пошук залишку (розділ ІІ);

 б) пошук залишку в обраній таблиці (розділ ІІІ);

 в) вибір потрібного суфікса з декількох омонімічних або спряжених – звісно, якщо знайдений суфікс має такі (розділ ІV);

 г) перенесення інформації від суфікса, що знайдено, до інформації до словоформи (розділ V);

 д) вибір подальших дій залежно від властивостей суфікса, що оброблено (розділ VI).

1. “Особлива” обробка вводиться заради вказаних вище (см. п.2) ускладнень, що притаманні “особливим” випадкам. Шляхом послідовних перевірок встановлюється, яке саме ускладнення має місце, і залежно від цього виконується одна з дій:

 а) одну з декількох омонімічних (або спряжених) основ, що її вибрано невірно для даної словоформи, замінюють на іншу (розділ VII);

 б) відтинається чергова основа складного слова (розділ VIII);

 в) розбираються випадки вкладення суфіксів у наступну основу складного слова (розділ IX);

 г) графічні чергування зводяться до прийнятої форми (розділ Х).

# Частина Б – “Список суфіксів”

 Частина Б являє собою список суфіксів, оформлений у вигляді однієї або кількох таблиць.

 Вирішення питання про те, розміщувати суфікси в одній чи в кількох таблицях, залежить від властивостей певної мови. Точний характер цієї залежності поки що невідомий, а тому доводиться керуватися загальними приблизними міркуваннями:

Якщо загальна кількість суфіксів у даній мові невелика (наприклад, близько двох десятків в англійській мові), їх можна об’єднувати в одну таблицю. Якщо ж загальна кількість суфіксів у даній мові велика (більше ста п`ятдесяти в українській мові), то є сенс розташувати їх в окремих таблицях відповідно до частини мови; при цьому звертання до тієї або іншої таблиці визначається ознакою частини мови у основи.

В мовах можуть зустрітися такі омонімічні суфікси, що вибір потрібного суфікса зумовлюється лише його позицією відносно інших суфіксів у залишку.

Наприклад, вибір одного з декількох омонімічних суфіксів α залежить від того, чи був даний суфікс першим у ланцюжку суфіксів, або йому передував один із суфіксів β, γ, δ, ε, ζ і т.д.

Щоб враховувати такі складні взаємозалежності суфіксів за можливістю стандартним способом, можна розподіляти подібні омонімічні суфікси по кільком таблицям таким чином, щоб у кожній таблиці було наведено лише один з омонімічних суфіксів. Окрім того, під час розподілу суфіксів по таблицях враховується їх взаємна сполучність. Нехай, наприклад, в наведеній вище ілюстрації суфікс α має значення *х*, коли він починає залишок або коли йому передує суфікс ε; омонімічний йому суфікс α має значення *у*, коли він знаходиться після суфікса β, γ або δ; нарешті, ще один омонімічний суфікс α має значення *z*, коли йому передує суфікс ζ. Тоді суфікси α, β, γ, δ, ε та ζ можна розташувати по таблицях таким чином:

у таблиці 1 ­– ε;

у таблиці 2 – α, β, γ, δ (у значенні *х*);

у таблиці 3 – ζ;

у таблиці 4 – α (в значенні *у*);

у таблиці 5 – α (в значенні *z*).

Послідовність пошуку залишків по таблицях така: таблиця 1 – таблиця 2 – таблиця 4 - таблиця 3 – таблиця 5. Те, в якій таблиці буде знайдено залишок, однозначно визначає значення суфікса (тобто, вибір потрібного з кількох омонімічних).

Такий підхід може бути зручним, наприклад, в алгоритмі морфологічного аналізу грузинської мови (для дієслівних закінчень).

Коли в частині Б суфікси розташовані в кількох таблицях, кожній таблиці дається номер тієї наступної таблиці, де потрібно шукати робочий залишок, якщо його не буде знайдено в поточній. Цей номер позначається через ψ і використовується в розділі ІІІ частини А.

# Частина В – “Список інформацій до суфіксів”

 Частина В – це перелік основних інформацій до суфіксів, що зібрані до однієї таблиці.

 Основна інформація до суфікса складається з низки інструкцій двох видів: 1) до яких рядків інформації до робочої словоформи надходять нові дані від суфікса, що знайдено; 2) які ці нові дані, тобто які нові значення повинні бути записані до вказаних рядків робочої інформації.

#

# Частина Г – “Нестандартний запис”

 Ця частина являє собою набір правил, що забезпечують розбір випадків, коли на результатах аналізу словоформи відбиваються індивідуальні особливості якихось морфем. Всі такі випадки об`єднуються (доки це можливо) у типові, і для кожного типового випадку в частині Г є окремий розділ. Таким чином, кількість розділів частини Г теоретично необмежена і залежить від кількості типових особливостей морфем у даній мові. Для російської мови, наприклад, частина Г має дванадцять розділів.

 Кожний розділ – це невеличка інструкція, що складається з кількох операторів і вказує, що саме потрібно робити, щоб, будуючи загальну інформацію до словоформи, правильно врахувати індивідуальні особливості морфем, що її складають.

# Частина Д – “Омонімія словоформ”

 Частина Д необхідна для аналізу таких випадків, коли від двох різних основ можуть утворюватися однакові форми (при цьому інші форми цих основ не співпадають). Інакше кажучи, тут розглядається те, що традиційно називають “граматичною омонімією”. Очевидно, що якщо дві форми співпадають, то їх основи також або співпадають повністю, або одну з них спряжено іншою. Тому, якщо в інформації до основи є вказівка на те, що від цієї основи утворюються деякі форми, що співпадають з формами якоїсь іншої основи, то в цій інформації обов`язково повинна бути або адреса інформації до омонімічної основи, або адреса інформації до спряженої основи.

 Частина Д складається з двох розділів: алгоритму обробки омонімічних словоформ (Д-І) та таблиці омонімічних словоформ (Д-2).

## Висновки

 Завдяки сучасним методам морфологічного аналізу створюються все більш досконалі морфологічні процесори. Але треба пам’ятати про один важливий аспект:

 Спроби побудови алгоритмів автоматичного аналізу природної мови завжди зводяться до спроб формалізувати розумову діяльність людини. І незважаючи на те, що спроби ці підчас дуже вдалі, завжди знаходиться щось, що не піддається формалізації.

Це й відрізняє людину від машини.

## Література

1. И.А. Мельчук «Морфологический анализ при машинном переводе», стаття із збірки “Проблемы кибернетики”, М., 1961.
2. Л.Н. Иорданская «Морфологические типы основ русского языка (для различения омонимии морфем в процессе анализа при машинном переводе)», стаття із збірки “Проблемы кибернетики”, М., 1961.
3. «Использование ЭВМ в лингвистических исследованиях», отв. ред. В.И. Перебейнос, Киев, «Наукова Думка», 1990.