

АНАЛІЗ ЗАДАЧІ ПОБУДОВИ ОНТОЛОГІЧНОГО СЛОВНИКА БУДІВНИЦТВА

Розглядається проблема автоматизованої побудови онтологій в будівництві на основі семантичного аналізу текстів природною мовою. Найважливішою проблемою в розвитку досліджень в будівництві є його інтелектуалізація, і пов'язані з цим інтеграція даних та якісний пошук інформації. Однак дані технології припускають наявність якісних джерел семантичних даних. В цілому, можна стверджувати, що повторне використання та інтеграція з джерелами даних знаходиться на низькому рівні. Це пов'язано з тим, що джерела даних не є досить досконалими, щоб їх можна було із зручністю використовувати або інтегрувати. У зв'язку з цією проблемою, завдання автоматизованого формування онтологій на основі аналізу текстів природною мовою в будівництві є досить актуальним. Поширений критерій якості онтології заснований на оцінці зручності та ефективності роботи з нею та наявності зв'язку когнітивних процесів з семантичними для її побудови і актуалізації в сфері будівництва. Оцінка побудованих онтологій на основі автоматизованого інструментарію є ще одним окремим складним завданням. Основною відмінністю проведеного дослідження, результати якого наводяться в даній роботі, є когнітивно-семантичний аналіз на основі теорії категорій, математичної логіки і універсальної алгебри, алгебри множин, алгебри кортежів і реляційної алгебри, а саме – побудова онтологічного словника будівництва, онтологічних конструкцій, відповідних тексту на відкритих мовах, подання знань в будівництві. В роботі проведена класифікація та визначення інформаційних одиниць та їх джерел. Визначені особливості інформаційного ресурсу в будівельній галузі.

Ключові слова: будівництво; онтологія; поняття; тезаурус; інформаційний об'єкт; ізоморфність; аналіз; семантична модель.

Актуальність і огляд досліджень. У будівництві в даний час спостерігається зростання об'єму інформації, що обумовлено збільшенням кількості інформаційних джерел, збільшенням ступеня її доступності, що призвело до втрати її систематизованості в цілому. Пошук інформації виконується за допомогою каталогів бібліотек, архівів і системи посилань, ключових слів, повнотекстової індексації, пошуку по тексту. В процесі будівництва можна виділити два типи суб'єктів:

- суб'єкти, які беруть участь в явній формі, а саме безпосередньо виконують будівельні операції на всіх етапах;

- суб'єкти, які беруть участь в неявній формі, безпосередньо здійснюючи операції забезпечення і управління, і таких суб'єктів значне число.

Для їх успішної взаємодії, на сьогодні, відсутнє формалізоване уявлення природних мов, і як результат, виникають протиріччя в поняттях: наявність в онтологічному словнику (ОС) поняття і його заперечення, вираженого в явній і неявній формі; введення в ОС нового поняття, що збігається за змістом з визначеним раніше; використання невизначених понять; протиріччя між окремими частинами поняття; протиріччя в предикатах - неправильне використання відносин при використанні.

Отримані результати в формалізації розуміння сенсу природних мов досягнуті на шляху використання аналогії між математичними і природничими мовами. Алгебро-логічний підхід до розуміння смислів текстів на природній мові розроблений в теорії інтелекту [1, 20, 26], що робить можливим вирішення завдань вилучення певної інформації – визначень і зв'язків термінів, іменованих сутностей на основі розпізнавання мовних конструкцій. Існує проблема стандартизації, уніфікації, нормалізації і інвентаризації понять, орієнтованих на фахівців в різних областях будівництва, які були б актуальні, поінформовані, багатомовні і допускали автоматизацію обробки науково-технічної інформації [11, 22]. З іншого боку, в рамках інформаційних технологій будівництва отримало подальшого розвитку дослідження стосовно інформації і моделювання в будівництві, інформаційного середовища в будівництві та його специфіка, аналіз моделей в інформаційному просторі будівництва, топологія інформаційного простору у будівництві, поле завдань об'єкта будівництва [2-7].

Актуальність створення онтологічного словника полягає в забезпеченні передачі користувачам гранично великого обсягу знань і досвіду, створення єдиного розподіленого інформаційного середовища в рамках предметної області (ПО) будівництва на основі інформаційних технологій будівництва.

Постановка задачі. У будівництві на основі дослідження співвідношення когнітивних і семантичних процесів *визначити модель утворення нових понять* в контексті реальності й умовності, згідно подій в залежності від онтологічного статусу в ПО для забезпечення інваріантного змісту поняття, визначеного онтологічно і гносеологічно, з метою побудови математичних моделей, що формалізують протиріччя, які можна виявляти і усувати автоматично на основі автоматизовано отриманої онтології репрезентативного корпусу мови ПО, які можна представити як елемент інтелектуальної системи.

Результати аналізу. Будівництво як діяльність є джерелом ОС. Інформаційний ресурс будівництва – це знання, які стали інформацією,

представленою як бази знань, декларативні знання, знання про предметну область в будівництві, механізми вилучення і придбання знань, процедурні знання, евристика і евристичні знання, експерти та експертні знання.

Під знаннями будемо розуміти вірне, апробоване практикою відображення реальності у формі інформації, що поєднана з цінностями людської організації, здатної при необхідності використовувати наявні дані відповідно до встановлених правил і процедур з урахуванням їхнього ставлення до цієї інформації. Таким чином, знання як результат процесів екстракції інформації є більш стійкими інформаційними утвореннями і розташовуються на більш високому рівні узагальнення, ніж дані і факти, що становлять інформацію, зокрема, знання відіграють роль структур, які забезпечують організацію даних і фактів.

Для знань характерні внутрішнє інтерпретування, структурованість, зв'язаність і взаємна активність.

Знання класифікуються на:

- декларативні знання – безпосередньо доступні для використання. Декларативні знання, як правило, використовуються для подання інформація про властивості і факти предметної області будівництва;
- процедурні знання – процедури, за допомогою яких можна отримати знання. Процедурні знання це інструкції, методики, інформація про способи вирішення завдань предметної області будівництва;
- евристичні знання – знання, що накопичуються інтелектуальним середовищем будівництва в процесі його функціонування;
- експертні знання – знання фахівців в деякій предметній області будівництва.

Необхідно особливо відзначити, в інтелектуальному середовищі на множині знань існує функція їх важливості по відношенню до суб'єкта предметної області будівництва. Це ілюструється, наприклад, розподілом знань на факти і відомості.

Факт – це знання у формі твердження, достовірність якого строго встановлена, припущення, що фіксує емпіричне знання будівництва.

Відомості – частина знань, критерій істинності яких не однаковий для різних учасників процесу будівництва їх утворення.

Для інтелектуальних систем характерні два процеси:

- вилучення знань – процес взаємодії суб'єкта з джерелом знань, в результаті якого стають явними структура і зміст уявлень предметної області будівництва.
- придбання знань – процес автоматизованого навчання шляхом передачі бази знань за допомогою діалогу.

Таким чином, знання інтелектуальної системи будівництва, яка породжується даним суб'єктом щодо даної предметної області складаються з наступних частин:

1) знання про предметну область будівництва – сукупність знань про предметну область, що зберігаються в базі знань інтелектуальної системи. Знання про предметну область поділяються на:

- факти, які стосуються предметної області будівництва;
- закономірності, характерні для предметної області будівництва;
- гіпотези про можливі зв'язки між явищами, процесами і фактами;

2) знання про суб'єкт предметної області будівництва – сукупність знань про суб'єкт предметної області, що зберігаються в базі знань інтелектуальної системи. Знання про суб'єкт предметної області поділяються на:

- процедури для вирішення типових задач у даній галузі;
- прийоми і методи прийняття рішень, що використовують інтуїцію і досвід фахівців у вирішенні проблем конкретної предметної області будівництва;
- експертні знання фахівців, які за роки навчання і практичної діяльності навчилися ефективно вирішувати завдання, які стосуються конкретної предметної області будівництва.

3) знання про інтелектуальну систему даного суб'єкта даної предметної області будівництва – сукупність знань про інтелектуальні елементи цієї системи як окремо, так і системні знання в цілому.

Як особливість інтелектуальних систем будівництва, відзначимо наявність у них специфічних знань – абстракцій і гіпотез.

Абстракція – одна зі сторін, форм пізнання, що полягає в уявному відстороненні від ряду властивостей предметів і відношень між ними і з виділенням, відокремленням якої-небудь властивості і відношення.

Гіпотеза – структурний елемент знань, що використовується в методі розвитку наукового знання, який включає в себе висунення і подальшу експериментальну перевірку припущень.

Знання в рамках ІСПОБ конгломеруються в таких структурах як семантична мережа та база знань.

Семантична мережа – структура даних, які отримуються за рахунок спостережень і навичок, що складається з вузлів, відповідних понять, і зв'язків, які вказують на взаємозв'язок між вузлами. Найбільш важливими зв'язками є зв'язки "Це-є", що дозволяють побудувати в семантичній мережі ієрархію понять, в якій вузли низьких рівнів успадковують властивості вузлів більш високих рівнів.

База знань – семантична модель, що описує предметну область будівництва і дозволяє відповідати на такі питання з цієї предметної області, відповіді на які в явному вигляді не присутні в базі. База знань є основним компонентом інтелектуальних систем будівництва. Система управління базами знань – це комплекс програмних, мовних і інтелектуальних засобів, за допомогою якого реалізується створення і виконання бази знань.

Узагальненою предметною областю будівництва даного суб'єкта будемо називати предметну область у взаємодії з даним суб'єктом.

Узагальненим суб'єктом даної предметної області будівництва будемо називати сукупність суб'єктів даної предметної області в їх взаємозв'язку.

Узагальнений суб'єкт даної узагальненої (універсальної) предметної області будівництва – це узагальнений суб'єкт, предметна область якого є універсальною для узагальненої предметної області кожного суб'єкта.

Наведена вище система понять і визначень дозволяє сформулювати визначення ІСПОБ (Інтелектуального середовища предметної області будівництва) як інформаційне середовище будівництва, що включає і породжує інтелектуальні елементи.

Таким чином, в рамках ІСПОБ утворення знання проходять всі етапи еволюції від виникнення, сприйняття і передачі знань "від людини до людини" до дифузії в інші ІСПО.

Накопичення знань в такій складній структурі, якою є ІСПОБ, неминуче повинно супроводжуватися ускладненням системи управління знаннями в цілому.

Особливістю утворення ІСПОБ є необхідність розрахунку п'яти послідовно зростаючих показників:

1) розрізнення (розпізнання) – рівень ознайомлення з процесом, об'єктом або явищем предметної області будівництва;

2) запам'ятовування – показник кількості засвоєної інформації, правил, формулювань;

3) розуміння – це усвідомлене відтворення як інформації в цілому, так і її складових частин;

4) застосування – вміння і навички застосування теоретичних знань на практиці будівництва;

5) розвиток – творчий рівень реалізації взаємоузгодження для різних завдань засвоєного теоретичного матеріалу і практичних навичок.

На поточний момент, знання будівництва, як такі, набувають все більш ресурсний характер. Це пояснюється тим, що поступово виснажуються природні ресурси і ресурси середовища проживання людини, тому в своїй діяльності вона змушена використовувати накопичені знання і досвід для більш глибокої переробки природних ресурсів. Ресурсність знань будівництва укладена в їх визначенні, як сукупності інформації (дані і факти) з їх цінністю для інтелектуальних середовищ будівництва, тобто знання це структурована інформація ІСПОБ в сукупності з оцінкою її значущості для даного ІСПОБ.

Визначимо знання як $Z \equiv \{U, S, I, K(U, S, I)\}$, де:

Z – знання в будівництві; U – інформація в ІСПОБ(I);

S – структура інформації будівництва;

K – значущість інформації будівництва;

I – інформаційне середовище будівництва.

Специфіка знань як самостійного ресурсу, який слабо заміняється іншими видами ресурсів, визначається тим що існує зв'язок знань з їх носіями і, як наслідок, в ІСПОБ необхідне існування механізму відділення знань від носіїв.

Введемо позначення:

$N = \text{support } I$ – множина всіх носіїв знань інформації з I ;

$n_i \in N$ – i -й носій знань на основі інформації з I ;

$J(I)$ – знання на основі інформації з I .

Визначимо властивості знань на основі інформації:

1. $\exists j \in J(I): Z_j(n_i) = Z_{ji} \neq \emptyset$ – породжене ІСПОБ носієм n_i щодо знань з індексом j ;

2. $Z_j^1 = \bigcup_i \text{support } Z_{ji}$, Z_{ji} – i -й носій знань j в ІСПОБ;

3. $\forall j \in I, \forall i \in N \exists \varphi_i^j : Z \rightarrow \{0,1\}$, де $\varphi_i^j \left(n_i = \begin{cases} 1, & Z_{ji} \in Z_j^1, \forall i \\ 0, & \exists i : Z_{ji}(n_i), Z_{ji} \notin Z_j^1 \end{cases} \right)$,

$\Phi = \{\varphi_i^j\}$ – оператор відділення знань від носіїв, тобто $\text{support } Z_j(\Phi(n_i)) = \text{support } I$ при цьому $\Phi = \bigcup_k \Phi(t_k, I)$, t_k – час.

У ІСПОБ відділення знань від носіїв здійснюється двома способами:

1) в напрямку від системи, що навчається, до навчального матеріалу, що навчається за рахунок подання знань, у вигляді НСІ;

2) у напрямку від того, хто навчається, до системи, яка навчається, за рахунок синтезу нового інтелектуального елемента на базі того, хто навчається, з подальшою адаптацією до ІСПОБ на основі уточнюючих питань, контрольних питань і запитів додаткової інформації.

До специфіки знань як ресурсу будівництва відноситься факт відсутності закону збереження цінності інформації, як складової знання, при передачі її іншому носію знань. Тому в умовах ІСПОБ функція контролю знань у того, хто навчається, реалізує критерій, який повинен бути, з одного боку, максимально об'єктивним, а з іншого – може чисельно обчислюватись.

Як наслідок, існує передачі знань від системи знань як носія до носія як інтелектуального елемента будівництва, який навчає, а саме:

$$\forall j \in J, \forall i, k \in N \exists \psi_{ik}^j : n_i \rightarrow n_k : \varphi(\psi_{ik}^j) = \varphi_i^j \cdot \varphi_k^j \neq 0, \Psi_{ik}^j = \{\psi_{ik}^j\}_{ik}.$$

Розглянемо інформаційні процеси в ІСПОБ:

- ініціація первинних знань: $t_0 = 0$, $\Phi(0, I) = 0$; $\forall t_1 > t_0$, $\Phi(t_1, I) \neq 0$, що відповідає виділенню підмножини в $\text{support } I$;
- переведення знань СБ, в систему знань: $\exists k_* \forall k_1 > k_* \forall j : \Phi_j(t_{k_1}, I) = \Phi_j(t_{k_*}, I)$;

- втрата знань в ІСПОБ I – виключення множини з підмножин $\text{support } I$ в момент часу $t_k \in T$.

Тоді: $(\Phi, \Psi_{ik}^j, I, Z, T)$ – процес поширення знань на I в ІСПОБ.

Визначимо цінність знання як дихотомічну функцію «може-не може» щодо носія знання і його дії-наміру в даний момент часу:

$$P(n) = \sum_{m=1}^M P_m(I, j, n, t_k, d(n)) \equiv \begin{cases} 1, & Z_{ji}(t_k) = 1 \\ 0, & \text{інакше} \end{cases},$$

$d(n)$ – сукупність інформаційних одиниць, необхідних суб'єкту для вирішення завдань в поточний момент часу t_k ; M – кількість задач що вирішуються; $P(n)$ – чисельно відповідає кількості завдань, що розв'язуються носієм n , в даний момент часу в даній предметній області будівництва і, що дозволяє ввести функцію вартості знань:

$$S_t(j, n) = \begin{cases} 0, & Z_{ij} = 0 \\ \sum \text{витрат ІСПОБ}, & Z_{ij} \neq 0. \end{cases}$$

ІСПОБ породжує поняття за рахунок формалізації опису та уніфікації уявлення виділеного об'єкта середовища будівництва. Уніфікація реалізується за рахунок включення в ІСПОБ нових понять, визначень і процедур на рівні онтології будівництва на її рівні уніфікуючих процедур. Формалізація опису, як процесу, базується на власне описі об'єкта середовища будівництва як функція F_f так і на класифікаторі термінів і понять прийнятому на даному інтелектуальному середовищі будівництва Cl .

Спочатку, джерела знань ІСПОБ є результатами досліджень різних авторів або колективів авторів. ІСПОБ породжує інтелектуальні системи, тому воно є локальним щодо його елементів, класифікаторів термінів і понять, а також, функції-формалізатори можуть відрізнятися, не бути тотожними або не збігатися:

$$Cl(m_i) \neq Cl(m_j), F_f(m_i) \neq F_f(m_j),$$

де $m_i \neq m_j$ – різні інтелектуальні елементи.

Нехай M – кількість суб'єктів будівництва, яким передаються знання, $M \cap N \neq \emptyset$.

Дослідник здійснює діяльність засобами і за рахунок своєї внутрішньої мови, використовуючи внутрішні знання ІСПОБ. Транслює синтезує ним поняття в терміни доступні будь-якому користувачеві даного ІСПОБ, що навчається, за рахунок прийнятих загальних класифікаторів термінів, понять (Cl) і функцій-формалізаторів (F_f).

У термінах підходу формальних мов, функціонування ІСПОБ на етапі синтезу понять еквівалентно реліктовому формалізатору будівництва, а в подальшому, транслятору в терміни сучасної мови

досліджень в рамках даної ІСПОБ. При цьому, трансляція розглядається як визначення нових понять з урахуванням їх відношення до досліджуваних об'єктів в ІСПОБ, в тому числі і середовищ, тобто їх класифікація та сприйняття ІСПОБ інформації про них є знанням щодо ІСПОБ.

Особливо необхідно відзначити наступні властивості:

$$Cl(I) \subseteq Cl(m_j) \forall i \in M, \quad F_f(I) \subseteq F_f(m_j) \forall i \in M.$$

Відносно Cl та F_f існує множина операцій:

$$Ved(\bullet) \stackrel{<def>}{\Leftrightarrow} \{ins(\bullet, \bullet), del(\bullet, \bullet), cor(\bullet, \bullet)\}, \text{ де:}$$

$ins(A, a)$ –включити (приєднати) елемент (об'єкт) a в A ;

$del(A, a)$ – виключити елемент (об'єкт) a з A ;

$cor(A, a)$ – коригувати (змінити) елемент (об'єкт) a в A .

Так як елемент $Ved(\bullet)$ класифікований в $Cl(\bullet)$ і описується за допомогою $F_f(\bullet)$, то для I справедливо: $Ved(m_i) \subset Ved(I), \forall i \in M$.

Так як для $i \neq j, Cl(m_i) \neq Cl(m_j)$, то ІСПОБ I є середовищем для утворення нових понять на їх базі знань, які в подальшому за рахунок процедур класу Φ стануть значеннями ІСПОБ, тобто:

$$J = J(m_i), \quad \bigcup_i \bigcup_{j \in j(m_i)} Z_j(m_i) \supset Z(I), \quad Z_{int}(I) = \bigcup_i \bigcup_{j \in j(m_i)} Z_j(m_i) \setminus Z(I),$$

$$\exists t_{k_0} \text{ support} \Phi(t_{k_0}, I)[(Z_{int}(I))] = \text{support} \Phi(t_k, I)[(Z_{int}(I))].$$

Ця рівність виражає суть того факту, що в будь-який момент часу внутрішні знання ІСПОБ є скінченними і можуть бути перетворені в ресурс.

Слід зазначити, що істинність або хибність цього твердження щодо знань отриманих із зовнішнього середовища будівництва, тобто за рахунок їх адсорбції ІСПОБ відповідно до своїх уявлень і на основі внутрішніх класифікаторів, відповідно до теореми Геделя встановити неможливо – система не повна, то цю властивість можна взяти як критерій відмінності ІСПОБ від будь-якого середовища як такого, тобто не обов'язково інтелектуального.

Для ІСПОБ справедлива наступна **теорема**: для ІСПОБ вираз $\exists t_{k_0} \text{ support} \Phi(t_{k_0}, I)[(Z_{int}(I))] = \text{support} \Phi(t_k, I)[(Z_{int}(I))]$ завжди є справедливим, у той час, як для довільних середовищ істинність виразу неможливо встановити. Цей факт підтверджує те, що розвиток мов завжди супроводжується науковим розвитком і розвиток науки призводить до розвитку мов.

Наслідок 1. $|J(I)|$ – є не спадною у часі функцією. Тобто, спостерігається збільшення концентрації знань у даній галузі будівництва з плином часу.

Наслідок 2. $|Cl(I)|$ – є не спадною у часі функцією. Тобто, споживши внутрішній ресурс $Z_{int}(I)$ оновлення класифікатора $Cl(I)$ буде здійснюватися за рахунок знань понять середовища предметної галузі будівництва у взаємозв'язку з суміжними предметними областями (дифузія інформації).

Характерним процесом, пов'язаним зі знанням як ресурсом є процес еволюції знань в інтелектуальних середовищах будівництва, що проявляється як в кількісному так і в якісному аспекті знань. Для оцінки кількісного зростання функцій $|J(I)|$, $|Cl(I)|$ на практиці використовується експоненціальна модель зростання словників, що враховує такі фактори як семантичну інтерференцію мов, архаїзми, новизми, ідіоми і синоніми. Специфікою еволюційного процесу знань є пере відкриття інформаційних констант, отримання еквівалентних інформаційних констант з подальшим їх заміщенням новими загальними знаннями, при цьому еквівалентність знань в рамках однієї інтелектуальної системи будівництва розуміється традиційно – як мовна еквівалентність понять.

Наступною специфічною рисою еволюції знань будівництва є наявність управління знаннями як ресурсом за рахунок продукування інтелектуальної складової середовища будівництва.

Визначимо особливості інформаційного ресурсу будівництва.

1. Інформація як категорія будівництва різна за своїм прагматичним призначенням:

- a) змістовно-фактуальна (ЗФІ);
- b) змістовно-концептуальна (ЗКІ);
- c) змістовно-підтекстова (ЗПІ);

і складається з наступних розрядів лексичних одиниць:

- a) слова природної мови в значенні, яке в ній прийняте;
- b) слова природної мови, які в професійному обороті будівництва вживаються, як правило, у вузькому, спеціальному значенні;
- c) слова, характерні для наукових текстів і лише зрідка вживані в ненаукових текстах як запозичення.
- d) фразеологічні вирази будівництва;
- e) спеціальна термінологія будівництва;
- f) слова загальнолітературної мови;
- g) символи, гістограми, схеми, математичні вирази і інші види несловесної наукової семіотики.

2. Інформація це інтелектуальний ресурс, фактор колективної творчості.

3. У міру споживання знань будівництва її обсяг не зменшуються.

4. По мірі використання обсяг інформації не зникає, а зберігається і навіть збільшується.

5. Інформація це не самостійний ресурс будівництва і сам по собі має лише потенційне значення.

6. Є формою безпосереднього включення науки до складу будівництва.

7. Інформація виникає в результаті не просто розумової праці, а його творчої частини.
8. Будівництву, як виду діяльності, притаманний суб'єкт будівництва (СБ), що є її інтелектуальним елементом, який реалізує когнітивну функцію і має наступні властивості:
 - a) різні СБ мають потенційно однакові канали отримання інформації;
 - b) в рамках фіксованого каналу отримання суб'єктивної інформації, ця інформація про один і той самий об'єкт для різних СБ в основному збігається (тобто СБ притаманне адекватне сприйняття середовища);
 - c) між різними СБ існують вербальний, знаково-символьний, образний комунікаційні канали;
 - d) для кожного СБ по відношенню до властивостей предметів або середовищ існує критерій «прийнятно - не прийнятно - байдуже»;
 - e) кожен СБ може створювати свою предметну область в будь-якому середовищі;
 - f) кожен СБ асоційований зі своїм середовищем;
 - g) кожен СБ має свої внутрішні уявлення (відображення) про середовище і предмет будівництва та їх властивості, які можуть бути їм передані за допомогою одного з доступних комунікаційних каналів;
 - h) кожен СБ може накопичувати, зберігати, передавати суб'єктивну інформацію;
 - i) кожен СБ може формулювати мету своєї діяльності, визначати механізм її досягнення, здійснювати отримання інформації і давати відповідні оцінки;
 - j) у структурі мовної особистості СБ присутні три плани:
 - структурно-семантичний, що відображає ступінь володіння повсякденною мовою;
 - лінгво-когнітивний, або тезаурусний (відображення ієрархії смислів і цінностей в тезаурусі);
 - мотиваційний (відображення в комунікації і сприйнятті сторонньої інформації, мотивів і цілей).

Виділимо особливості інформаційного середовища будівництва.

1. Множинність онтологій будівництва, як специфікацій концептуалізації предметних областей, з різних дисциплін.
2. Множинність словників користувачів онтологій будівництва.
3. Множинність суб'єктів будівництва і їх варіабельність.
4. Недосконалість середовища інформаційного обміну будівництва.
5. Наявність будівельних стандартів.
6. Різноманіття виробників для однотипних МТР (матеріально технічних ресурсів).
7. Існування сукупності технологій проектування та будівництва.
8. Множинність інформаційних джерел і інтелектуальних елементів в середовищі будівництва.

9. Існування синонімії для ідентифікаторів позначення об'єктів і суперечливої інформації в середовищі будівництва.
10. Онтологію проблематики будівництва – як семантичному полі модальної предметної області.
11. Наявність ознак, властивостей, понять, обумовленості, нечіткості в середовищі предметної області будівництва.
12. Залежність ідентифікатора інформаційного об'єкта будівництва від розв'язуваної задачі (він може вказувати на різні властивості об'єктів).
13. Існування відношення «замінимість - застосовність» над властивостями МТР в залежності від розв'язуваних завдань і суб'єктів будівництва.
14. Існування відношень між ідентифікаторами (атрибутами) і поняттями (денотатами) в середовищі будівництва, що мають такі властивості:
 - a) ідентифікатор може бути не пов'язаний з жодним поняттям;
 - b) ідентифікатор може бути пов'язаний з багатьма поняттями і є для них атрибутом;
 - c) ідентифікатор може бути пов'язаний не з одним поняттям і є для нього атрибутом;
 - d) ідентифікатор може бути пов'язаний з багатьма поняттями і не є для них атрибутом, а вказує на якісні властивості.
15. Синергетичність [27] – джерело породження понять на множині властивостей, ознак і завдань в середовищі будівництва.
16. Достовірність та актуальність даних в інтелектуальному середовищі предметної області будівництва (ІСПОБ) є фундаментальними поняттями, зокрема, в більш широкому розумінні ніж в інформатиці. До об'єктивних причин, що призводить до зменшення достовірності даних в ІСПОБ, відносяться:
 - a) неповнота уявлень про об'єкти дослідження в ІСПОБ. що є фундаментальною для аксіоматичних систем до яких вона відноситься;
 - b) помилка вводу-виводу, що є неминучою в людино-машинних системах;
 - c) технічні збої, спотворення та втрата інформації, що зберігається в ІСПОБ, як фактор організаційно-технічних систем;
 - d) помилка передачі даних будівництва як фактор використання комунікаційних систем;
 - e) неприпустимі трансляції в системах ІСПОБ - ІСПОБ і необґрунтовані процедури виведення даних, включаючи недостовірні методики, як фактор діалогових систем;
 - f) інформаційна невизначеність (неоднозначності) будівництва – випадки, коли інформація має кілька розумінь (інтерпретацій);
 - g) невизначеності, що не усуваються контекстом, омонімії і полісемії лексичних одиниць і синтаксичних конструкцій. Під актуальністю даних, з точки зору ІСПОБ, розуміється їх здатність підвищити достовірність всієї сукупності даних з точки зору поточних цілей

будівництва в рамках тенденції розвитку ІСПОБ в цілому. Процедурою, яка істотно впливає на актуальність даних в ІСПОБ, є процедура їх глобалізації, зокрема, вона робить загальнодоступними дані, отримані окремими СБ як в рамках даної ІСПОБ, так і за рахунок трансляції їх з інших ІСПОБ;

h) наявність пасток в ІСПОБ, а саме:

- поставлене завдання не можна вирішити в ІСПОБ.
- завдання в рамках даної ІСПОБ для свого рішення вимагають не допустимого часового ресурсу;
- завдання в рамках даної ІСПОБ для своєї постановки вимагає не допустимого часового ресурсу;
- постановка завдання в цілому в рамках даного ІСПОБ принципово грубо апроксимує завдання, що вирішується;
- постановка завдання в цілому в рамках даного ІСПОБ грубо апроксимується, а вдосконалення цього ІСПОБ вимагає неприпустимих ресурсів;
- завдання має безліч рішень, що можуть допускати взаємне виключення;
- завдання має безліч постановок, у тому числі такі, що не приводяться одна до одної в термінах даного ІСПОБ;
- здійснення рішення однієї і тієї ж задачі тільки в різних термінах даного ІСПОБ – відсутність коректного механізму перевірки виведення;
- наявність і використання не виявлених в ІСПОБ помилок;
- наявність і використання в ІСПОБ слабо детермінованих або мало значущих знань;
- ускладнений висновок необхідних вихідних знань, або для їх отримання потрібно не допустимий часовий ресурс;
- множинність коригувань в постановці завдання – відсутність розуміння мети або навичок у вирішенні подібних завдань;
- постановка «порожніх» завдань – задач, які якісно відрізняються від реальних завдань в даній предметній області, але сформульованих в даній ІСПОБ, невміння виділити основні або ранжувати за важливістю наявні фактори;
- неможливість розширити ІСПОБ – відсутність розуміння про необхідні знання, відсутність розуміння механізму розширення, не врахування знань для вирішення задачі про своє ІСПОБ, тобто знань які означають, що ІСПОБ може аналізувати свої власні дії і тому в змозі пояснити, як і чому воно вирішує конкретні завдання;
- розширення ІСПОБ в цілому для вирішення завдання, вимагає неприпустимих ресурсів;
- необхідність постійного розвитку універсального – агрегуючого, дезагрегуючого, інтерпретатора понять і термінів, так як в

постановках і при вирішенні задач використовуються знання різного рівня агрегації і для їх вирішення необхідні нові методики;

- рішення «не тих» завдань, через неадекватність розуміння або тлумачення термінів в ІСПОБ користувачем;

- обмеженість класифікації користувачів, множинність різних профілів користувачів – як наслідок, аморфний і слабо спеціалізований діалоговий інтерфейс користувача;

- динаміка ступеня само ідентифікації ІСПОБ – необхідність множини налагоджень, налаштувань, інсталяцій засобів ІСПОБ.

17. Судження СБ в ІСПОБ як правило представлені в тризначних логіках – судження може бути або істинним, або хибним, або невизначеним. На множині суджень можуть бути задані відношення за значеннями істинності, які можна порівнювати (сумісні: еквівалентність, логічне підпорядкування, частковий збіг) і які не можна порівнювати (несумісні: протилежність, суперечність) і незрівнянні.

18. Алетичність модальності суджень будівництва – це об'єднання логічних і онтологічних модальностей, які включають такі модальні оператори або категорії модальності: необхідність і випадковість, можливість і неможливість.

19. Умовивід, що є формою абстрактного мислення будівництва в якому з одного або декількох суджень - посилянь на підставі певних правил виведення виводиться нове судження, з необхідністю або певним ступенем ймовірності. Умовивід ділиться на такі види: дедуктивні, індуктивні, за аналогією, дедукції правила прямого виведення і правила непрямого виведення.

20. Технології в будівництві представляють кумулятиви інтелектуальних знань, що включають наступні технології:

- a) архітектурні;
- b) проектування;
- c) власне будівництва;
- d) реставрації будівельних об'єктів;
- e) реконструкції будівельних об'єктів;
- f) управління будівництвом;
- g) будівельної геодезії;
- h) застосування будівельної техніки;
- i) застосування будівельного обладнання;
- j) застосування будівельних конструкцій;
- k) застосування будівельних матеріалів;
- l) управління персоналом в будівництві.

21. Структура концептуальної моделі предметної області будівництва включає:

- a) опис об'єктів, що представлені на природних мовах;

- b) опис понять і відношень об'єктів інформаційного простору будівництва (в тому числі семантичних, які описуються наборами значень властивостей);
- c) ідеографічний словник предметної області будівництва;
- d) сукупність відповідностей між мовними одиницями і елементами концептуальної моделі;
- e) таксономію властивостей інформаційних об'єктів, як алгебраїчну систему, що задає склад та взаємозв'язки властивостей понять предметної області – властивості відображають семантику даного об'єкта, оскільки одного імені для позначення сенсу об'єкта недостатньо в силу полісемії природної мови.

Структура концептуальної моделі предметної області будівництва передбачає наявність у своєму складі тезауруса - словника, із зазначенням семантичних відношень (синоніми, антоніми, пароніми, гіпоніми, гіпероніми) між лексичними одиницями, зібрання відомостей, корпус або звід, повномірно охоплюючі поняття, визначення і терміни будівництва, а також передбачає наявність у своєму складі онтології предметної області будівництва.

Під онтологією будівництва (ОБ) будемо мати на увазі явний опис:

- множини об'єктів і зв'язків між ними, що складається з понять, організованих в таксономії,
- визначень, правил виведення, що не обмежуються онтологіями наук, які формують будівництво і не зводяться до них і визначає зміст теорії будівництва,
- об'єктів, які постулюється як існуючі.

У даній роботі предметна область будівництва описується концептуальною моделлю і онтологією – формою, придатною для семантичного аналізу на рівні предметної області будівництва. Онтологія в вигляді семантичної мережі виступає, як об'єднання всіх семантичних уявлень інформації з корпусу предметної області будівництва в єдину мережу, що робить можливим її автоматичне отримання, якщо існує можливість опису у вигляді набору тестів на природній мові.

Семантичне ядро предметної області будівництва – це словник слів - понять, що мають істотне значення для виявлення і класифікації інформації, яка відноситься до предметної області будівництва і характеризується термінологічністю, формальністю визначень, полісемантикою і наявністю спеціалізованих словників. Такий словник будемо називати словник онтології будівництва (СО). Для актуалізації тієї чи іншої семантики в СО служить контекст. Сукупність усіх словникових статей утворює корпус словника СО [15,16,17,18,19].

Виділення семантичного ядра будівництва можливе за рахунок того, що:

- регламентована послідовність будівельних процесів проектування і реалізації проекту об'єкта будівництва;

- в рамках будівельних процесів визначено послідовність процедур будівництва;
- в рамках будівельних процедур визначена послідовність дій;
- визначено застосовність – замінимість будівельних матеріалів, комплектуючих і конструкцій;
- регламентовано застосування та режими використання будівельної техніки та механізмів;
- в будівельній геодезії регламентовані процедури і дисципліна їх застосування;
- у будівництві регламентований документообіг.

Локалізація місць і інтервалів появи нових понять будівництва здійснюється за рахунок визначення порушень порядку зазначених вище процесів та процедур з подальшою класифікацією порушень по суті.

До властивостей семантичного ядра будівництва відноситься повторюваність, стійкість структури, локалізація в інформаційному просторі, наявність синонімічних рядів для понять, що до нього входять.

Розглянемо поняття синонімії будівництва для дослідження механізму породження понять. Нехай задана множина об'єктів D_0 , підмножини D_1, D_2 $at \leftrightarrow dt$ (відповідність атрибутів ита денотатів), $\varphi_0, \varphi_1, \psi_0, \psi_1$ – процедури виводу. Одне поняття може бути отримано різними операціями з різних вихідних даних, а саме, з еквівалентності понять, які описуються різними термінами ($at_1 \neq at_2$). Розглянемо відношення на множині процедур виводу.

- Тотожність (\equiv) (рис. 1): $\varphi_0\varphi_1 = \psi_0\psi_1; \exists D_0, \exists \varphi_0, \exists \psi_0 : dt(E_1) = dt(E_2)$
 $(\equiv) \psi_0^{-1}(D_2) \cap \varphi_0^{-1}(D_1) \neq \emptyset$.

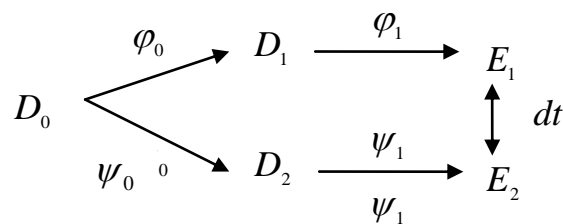


Рис. 1 Схема реалізації процедури тотожності

- Омонімія (рис. 2): $at(E_1) = at(E_2), dt(E_1) \neq dt(E_2), D_1 \cap D_2 = \emptyset$.

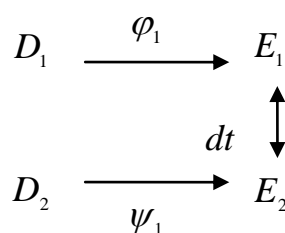


Рис. 2 Схема реалізації процедури омонімії

- Еквівалентність (\sim) (рис. 3, 4):

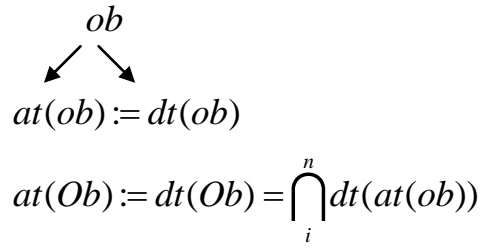


Рис. 3 Схема присвоєння атрибутів

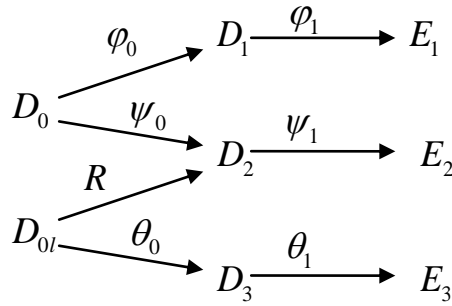


Рис. 4 Схема реалізації процедури еквівалентності

Визначені вище відношення мають наступні властивості:

1. Еквівалентність (\sim) породжує розбиття на класи, яким притаманна певна сутність відносно фіксованої множини;
2. Тотожність (\equiv) породжує розбиття на класи (рис. 5):

$$\varphi_1\varphi_0 = \psi_1\psi_0, \theta_1\theta_0 = \psi_1R \Rightarrow \psi_1 = \theta_1\theta_0R^{-1}, \varphi_1\varphi_0 = \theta_1\theta_0R^{-1}\psi_0$$

- Відповідність (C): $Ob \leftrightarrow at_1, dt \leftrightarrow at_2, at_1Cat_2$:

$$\left\{ \forall i \exists j \exists ob(i): at_1^i \leftrightarrow at_2^j \right\} \wedge \left\{ \forall j \exists i \exists ob(j): at_2^i \leftrightarrow at_1^j \right\}, \\
 \left\{ at_1^i, at_1^j \in at_1; at_2^i, at_2^j \in at_2; \forall ob \in Ob \right\} \wedge \left\{ \forall ob \in Ob \text{ властивість не виконується} \right\}$$

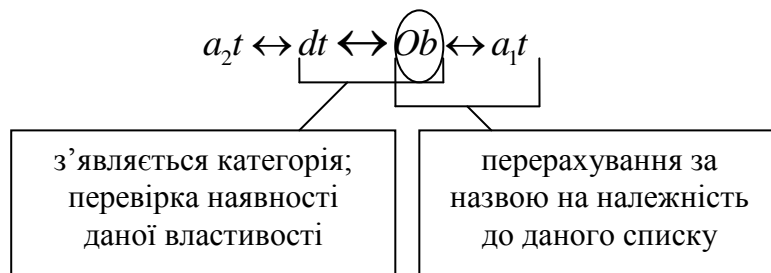


Рис. 5 Схема реалізації процедури відповідності

Процедура відповідності має наступні властивості:

1. C – рефлексивно, симетрично.
2. Об'єкти належать одному класу.
3. Ознаки належать одній категорії та $\{at_1Cat_2\} \wedge \{\overline{at_1}C\overline{at_2}\} \Rightarrow at_1 \equiv \overline{at_1}$ за рахунок перейменування атрибутів об'єктів, де $at_1, \overline{at_1}$ – множина атрибутів.

Для приведення до одного денотату визначимо синонімічні, еквівалентні та тотожні структури. Нехай:

$\{P_i\}$ – ознаки, $\forall i, j P_i \cap P_j = \emptyset$;

$B(\{P_i\})$ – структури (системи понять), B – буліан, $\{P_i\} \rightarrow B(\{P_i\})$;

$\forall i: S_i = \bigcup_{j=1}^{n_i} P_{ij}, \forall i: S_i \in B(\{P_i\})$.

Визначимо структуру понять як сукупність: $fix: \{S_i\}_{i=1}^n, S = \bigcup_{i=1}^n S_i$;

$S_0 = \bigcap_{i=1}^n S_i$, де: - $def: \{S_i\}_{i=1}^n$ еквівалентні по $S_0 \Leftrightarrow S_0 \neq \emptyset$;

- $def: \bar{S}$ покриття $\{P_i\}^{(\bar{S})} \Leftrightarrow \neg \exists P_*: P_* \notin \bar{S}, \forall i \neg \exists P_* \notin \{P_i\}^{(\bar{S})}$;

- $def: *(S_0) \xleftrightarrow{def} \bar{S} \setminus S_0$ – топологічна перед база;

- $\{S_i\}_{i=1, \infty}$ – еквівалентні по $S_0 \Leftrightarrow S_0 \subseteq \bigcap_{i=1}^{\infty} S_i S_0$ (рис. 6);

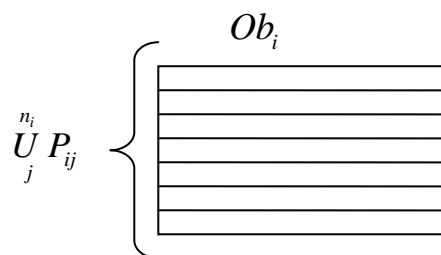


Рис. 6 Еквівалентність по S_0

- $\sum_{m=0}^{n_i} \frac{n_i!}{(n_i - m)!}$ – кількість ідентифікаторів, що виражені через ідентифікатори ознак;

- $C^{-1}(at)$ – синоніми: $\left\{ Ob \xrightarrow{C} At \right\} \wedge \left\{ \bigcup_{at \in At} C^{-1}(at) = Ob \right\}$ для скінченних множин Ob і At , $\dim At \geq \dim Ob$;

- $at_1 \neq at_2 \Rightarrow C$ – синтетична класифікація на Ob по At і C_{Ob}^{At} , $C^{-1}(at_1) \cap C^{-1}(at_2) = \emptyset$.

Визначимо представлення:

$$\left\{ C_{ob}^{at_1} \sim C_{ob}^{at_2} \right\} \xleftrightarrow{def} \begin{cases} \forall i \exists j \exists ob(i): at_1^i \leftarrow ob(i) \rightarrow at_2^j \\ \forall j \exists i \exists ob(j): at_2^j \leftarrow ob(j) \rightarrow at_1^i \\ |at_1| = |at_2| = |Ob| \end{cases}$$

Група перестановок на множині потужності $|Ob|$ задає всі синоніми на основі перерахування на множині $\{Ob\}$.

Наприклад, є синонімічність $S_0 = S_1 \cap S_2 \cap S_3 \neq \emptyset$ (рис. 7), тоді S_1, S_2, S_3 – еквівалентні відносно S_0 (рис. 8).

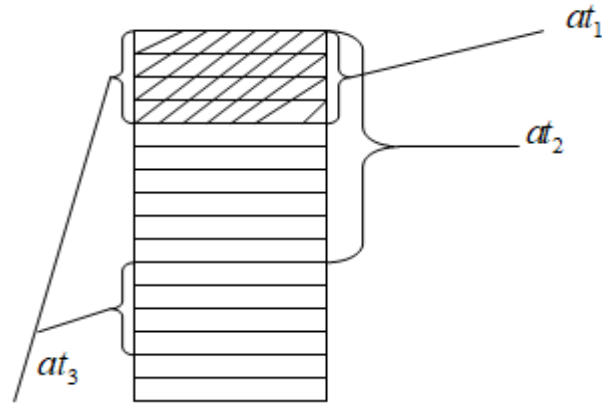


Рис. 7 Графічне представлення еквівалентності

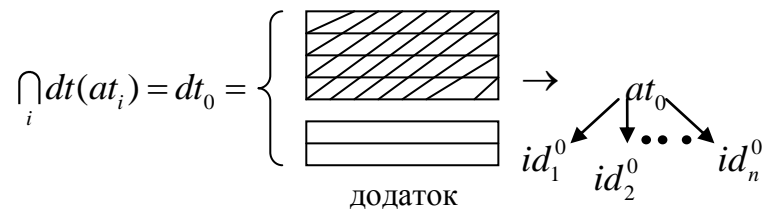


Рис. 8 Графічне представлення синонімічності

Нехай, $id_1^0, id_2^0, \dots, id_n^0$ – ідентифікатори синонімів. Тоді at_0 – коректно, якщо додаток відсутній або його не можна видалити по причині зв'язності (рис. 9).

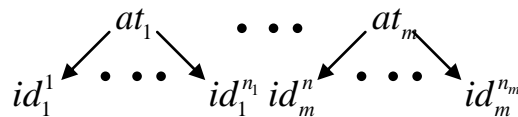


Рис. 9 Синоніми $at_1 \dots at_m$ по відношенню dt_0 на множині об'єктів Ob .

Для модуля $M = \{M_0, M_*, Op^M\}$ де
 $M_0 = \bigcup_i \bigcup_j L_{0j}^i$ – модуль вхідних початкових понять,
 $M_* = \bigcup_i \bigcup_j L_{*j}^i$ – модуль вихідних результуючих понять,
 $Op^M = \bigcup_i \bigcup_j Op_{ij}^M$ – визначення на модулі M ,
 $L = \{\{L_0^i, L_*^i\}, Op_i\}$ – процедури введення понять

справедливі наступні твердження:

$$\{\{L_{0j}^i \cap L_{*K}^i \neq \emptyset \Rightarrow K \geq j\} \cup \{L_{0j}^i \cap L_{*K}^i \neq \emptyset \Rightarrow \{i \leq m\}\}\}$$

$$\left\{ \left\{ L_{0j}^i \cap L_{*K}^i \neq \emptyset \Rightarrow Op_{ij}^M \cap Op_{iK}^M = \emptyset \right\} \cap \left\{ L_{0j}^i \cap L_{0K}^m \neq \emptyset \Rightarrow Op_{ij}^M \cap Op_{iK}^M = \emptyset \right\} \right\}.$$

Поняття в будівництві виникають в наступних випадках:

- при вирішенні однотипних завдань;
- при реалізації однотипних будівельних процесів;
- при практиці застосування однотипних методик в будівництві;
- при внесенні із зовні щодо будівництва понять, заснованих на нових знаннях;
- запозиченні з інших дисциплін «за аналогією», інновацій;
- виділенні інваріантної інформації;
- при веденні діалогу в межах ІСПОБ і формуванні поля завдань.

Спираючись на отримані раніше результати, проведемо дослідження діалогу, як елементу ІСПОБ. Користувач, що приступив до дослідження об'єкта (Ω), як правило, спочатку має про нього лише початкові уявлення (знання) $\{\omega_i\}$, або механізм, процедуру (Φ_1) отримання таких уявлень. У процесі вивчення об'єкта Ω у користувача складається модель об'єкта (M_1), яка відповідає початковим уявленням, а також механізм (процедура) (Ψ_1) синтезу спрощених моделей $\{m_i\}$ об'єкта Ω . Виключивши з розгляду так звані "моменти осяяння", у користувача в процесі дослідження має місце внутрішній діалог (D_1), спрямований на встановлення відповідності між його знаннями про об'єкт Ω і поданням M_1 даного об'єкта на рівні відповідних мисле образів. В силу особливостей конкретного користувача, в процесі умоглядного дослідження їм використовуються прості або спрощені моделі (W_1) об'єкта Ω з точки зору ІСПОБ, що представляють з цієї точки зору структури знань (S_1). Еволюція структур знань здійснюється двома шляхами:

- за рахунок дослідження уявлень про об'єкт:

$$S_1(\Omega) \longrightarrow W_1(M_1) \longrightarrow S_2(\Omega)$$

- за рахунок внесення нових знань:

$$S_1(\Omega) \longrightarrow S_2(\Omega)$$

В якості математичного апарату для вирішенні питань організації гнучкого діалогу, при якому народжуються нові поняття, може бути використаний апарат теорії графів діалогових процесів (ГДП) та операцій над ними. При такому підході виділяється набір базових ГДП, на основі яких стандартними операціями алгебри (процедурами СУБД) може бути синтезований комплексний ГДП, що відповідає задачі, яка розв'язується. Структура діалогу при заданому ГДП повністю визначається після опису сценарію діалогу, який задає набір дій, що виконуються в кожній вершині ГДП і необхідний контроль коректності синтезованих ГДП і сценарію.

Оскільки кожен дослідник об'єкта Ω представляє собою деяку структуру знань, і між дослідниками здійснюється обмін знаннями в рамках загальноприйнятих уявлень про об'єкт Ω на даний момент часу, що доповнюється переліком гіпотез (G) і переліком вирішуваних завдань (V), то діалог ІСПОБ допускає розгляд, як засіб організації обміну інформацією між ІСПОБ₁ і ІСПОБ₂, де ІСПОБ₂ доступна в даний момент часу досліднику (користувачеві) інтелектуальне середовище, що синтезується іншим користувачем. У термінах ІСПОБ описаний діалог можна представити у вигляді схеми (рис. 10),

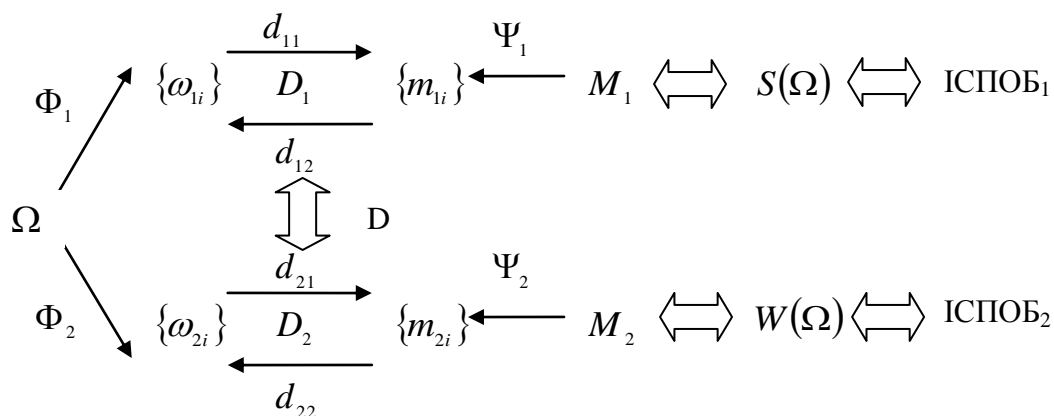


Рис. 10 Схема обміну інформацією між ІСПОБ

де під спряженим діалогом між дослідниками (або ІСПОБ): $D \Leftrightarrow D_1 \leftrightarrow D_2$ мається на увазі діалог, який допустив би існування і істинність схем, які представлені на рис. 11.

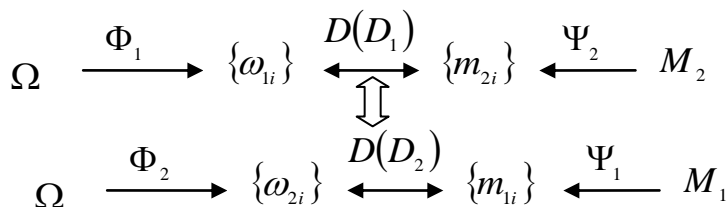


Рис. 11 Структура діалогу

Із структури діалогу випливає допустимість наступних еволюцій структури знань:

$$S_1(\Omega, M_1) \rightarrow W_1(\Omega, M_2) \rightarrow S_2(\Omega, M_2), S(\Omega, M_1) \rightarrow S_2(\Omega, M_2).$$

В основі еволюції лежить поява у користувача в даний момент нового уявлення (ω_0) про об'єкт Ω , або нового запиту (v_0), або нової гіпотези (g_0), що призводять в подальшому до трансформацій (рис. 12):

$$\begin{array}{lcl}
V(S) & \longrightarrow & V(S) \cup v_0 \\
V(W) & \longrightarrow & V(W) \cup v_0 \\
S(\Omega) & \longrightarrow & S(\Omega) \cup \omega_0 \\
G(W) & \longrightarrow & G(W) \cup g_0
\end{array}$$

Рис. 12 Трансформації появи нових знань

Перші три трансформації призводять до інтеграції, злиття або агрегації ІСПОБ.

Нехай:

$$W_1(\Omega_1) \equiv (\Omega_1, M_1, J_1, D_1, \tilde{\Phi}_1, \Phi_1^*), \quad W_2(\Omega_2) \equiv (\Omega_2, M_2, J_2, D_2, \tilde{\Phi}_2, \Phi_2^*),$$

тоді:

- інтегрування – $J_n(W_1, W_2) = W_3(\Omega, M, J, D, \tilde{\Phi}, \Phi^*)$ $M_2 \subset M_1, \Omega_1 = \Omega_2$;
- злиття – $Mg(W_1, W_2) = W_3(\Omega, M, J, D, \tilde{\Phi}, \Phi^*)$ за умови $\{M_2 \not\subset M_1\} \wedge \{M_1 \cap M_2 \neq \emptyset\}$ та $\Omega_1 \neq \Omega_2$;
- агрегація – $Ag(W_1, W_2) = W_3(\Omega, M, J, D, \tilde{\Phi}, \Phi^*)$ за умови $\{M_2 \not\subset M_1\} \wedge \{M_1 \cap M_2 = \emptyset\}$ та $\Omega_1 \neq \Omega_2$.

Розглянемо більш докладно процес внесення в модель W нової гіпотези g_0 .

Оскільки гіпотеза g_0 є елемент ІСПОБ, то існує запит v_0 в рамках ІСПОБ, що відповідає цій гіпотезі, точніше – якому відповідало б віртуальне уявлення ω_0 про об'єкт Ω . Так як гіпотеза g_0 є недоведеним твердженням, то це означає, що процедура верифікації (ідентифікації) або не реалізовувалася в ІСПОБ щодо g_0 , або не дає однозначної відповіді.

В останньому випадку можливі два варіанти:

- процедура не здійсненна через множинності уявлення m_{0i} , які можуть містити альтернативні затвердження;
- процедура не здійсненна через аксіоматичну неможливість виведення ω_0 , що призводить до збільшення розмірності аксіоматичного базису W .

Таким чином, розглянута трансформація відрізняється від попередніх трансформацій тим, що породжує або поповнює модель W віртуальним простором M , що, загалом, не призводить до еволюції ні моделі W , ні структури знань S , а є прототипом можливої еволюції в даному напрямку і одночасно при реалізації і верифікації (встановлення істинності) такої трансформації розглянуті раніше три трансформації будуть окремими випадками.

У ІСПОБ процедура верифікації здійснюється як ідентифікація. Звідки безпосередньо випливає: рішення задачі перевірки гіпотези може

бути здійснено за рахунок ресурсів іншого ІСПОБ, що відповідає поняттю, що вводиться.

У зв'язку з вище сказаним, дамо визначення узагальненого діалогового завдання (УДЗІСПО).

Визначення 1: Під УДЗІСПО мається на увазі синтез $g_0 \in M \sim$ для заданої $S(\Omega)$ в рамках обраного $W(\Omega)$.

Визначення 2: Під вирішенням УДЗ ІСПОБ мається на увазі відшукування або синтез $W_0(M_0)$ та спряженого діалогу $D: W(M) \rightarrow W_0(M_0)$ такого, що $g_0 \in M \sim \wedge \varphi_0 \sim$, де $\varphi_0 \sim$ – множина завдань $W_0(M_0)$, що вирішуються.

Для заданого об'єкта дослідження Ω розглянемо всі існуючі структури знань $\{S_i(\Omega)\}$ і моделі $\{W_j(\Omega)\}$. На базі $\{S_i(\Omega)\}$ відновимо відповідні їм моделі: $\{S_i(\Omega)\} \rightarrow \{W_{s_j}(\Omega)\}$.

Зв'яжемо всі моделі універсальним діалогом D в модель $W(\Omega)$ і в рамках її сформулюємо гіпотезу g_0 . Чи має така ОДЗІСПО рішення, залежить від істинності твердження g_0 в мові діалогу D . Тобто паралельно з побудовою моделі $W(\Omega)$, побудована модель, що відповідає схемі, представлений на рис. 13.

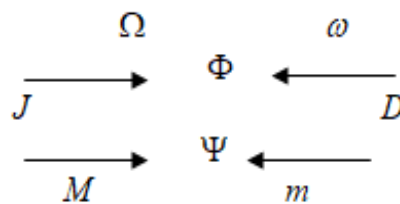


Рис. 13 Схема визначення приєднаної моделі

Модель, побудована за схемою (рис. 13) формально отримується з схеми ІСПОБ наступними замінами:

$$\begin{aligned} \Omega &\rightarrow J, & M &\rightarrow D, \\ \psi &\rightarrow m, & \Phi &\rightarrow \omega, \\ d_1 &\rightarrow \Phi, & d_2 &\rightarrow \Psi \end{aligned}$$

називається приєднаною до $W(\Omega)$, а відповідна УДЗ приєднаною до УДЗІСПО.

Доведена наступна теорема: УДЗ ІСПО і спряжена до неї УДЗ одночасно або мають рішення, або не мають.

Відзначимо, спряжена УДЗ ІСПО відповідає моделі функціонування об'єкта Ω , будучи його імітаційною моделлю.

Наведемо основні визначення необхідні для подальшого аналізу процесу появи поняття.

Відношення – математична структура, яка формально визначає властивості різних об'єктів і їх взаємозв'язків. Одномісні відношення відповідають властивостям. Для опису відношень застосовується апарат алгебри множин, алгебри Кодда, алгебри кортежів і реляційної алгебри.

Ознака – наявність або відсутність властивості у об'єкта, а також наявність або відсутність відношень між об'єктами. Ознаки, які обов'язково належать об'єкту, називаються істотними. Ознаки вибираються таким чином, що кожна ознака є необхідною, а в сукупності достатньою для поняття.

Поняття – відображена в мисленні єдність істотних властивостей і відношень предметів; думок, що виділяє і узагальнює предмети деякого класу за загальними і в своїй сукупності специфічними для них ознаками. В уточненому формулюванні: поняття є думка, яка за допомогою вказівки на деяку ознаку виділяє з універсуму і групує (узагальнює) предмети, що володіють цією ознакою.

Формальний **контекст** поняття є трійка: <множина об'єктів, множина ознак, ставлення інцидентності об'єктів і ознак>. Множина всіх понять формального контексту утворює повну решітку понять.

Властивість – атрибут об'єкта, прояв якості, є категорією, в булевій алгебрі і збігається з поняттям «предикат» – твердженням, висловленим про суб'єкт.

До факторів, від яких залежить процес утворення понять суб'єкту будівництва (СБ), відносяться:

- a) географія країни;
- b) традиції будівництва;
- c) тренди в будівництві;
- d) ступені комунікації з СБ інших країн;
- e) кваліметричні показники СБ;
- f) доступність і забезпеченість СБ матеріально-технічними ресурсами;
- g) нормативна база будівництва.

Система понять СБ складається з елементів трьох типів:

- a) поняття, запозичені з інших наук;
- b) оригінальні поняття і аксіоми;
- c) поняття низьких ієрархій, що розкривають зміст кожного з основних понять.

Формування понять в ОС відбувається при:

- a) класифікуванні або описі спостережуваних об'єктів в рамках існуючої онтології;
- b) дослідженні або встановленні відношень на поняттях онтології;
- c) запозиченні і об'єднанні онтологій різних предметних областей;
- d) трансляції під одним ім'ям, але в різних онтологіях різних об'єктів;
- e) приведенні онтології одного денотата (синоніми, синонімічні ряди – домінанта і інші уточнюючі члени ряду, які розширюють його семантичну структуру і доповнюють оціночними значеннями. Основна

причина виникнення синонімічних рядів – багатозначність терміна, а саме здатність багатозначних термінів входити в різні синонімічні ряди, інтерференцією необхідної однозначності);

f) виявленні нових властивостей у старих поняттях як окремо, так і емерджентно для їх груп;

g) необхідності уніфікації в поданні знань;

h) самоорганізації – виникненні нового порядку або структури в системі знань будівництва;

i) реконсиляції – перевірці цілісності прикладних даних в розподілених системах;

j) формуванні понять в ОС відбувається відповідно ланцюжку:

- зміст (зміст виростає з денотатів). Змістом поняття (інтенсіоналом) називається сукупність істотних ознак предмета, які можуть бути виділені в якості окремих понять.

- конотація (сміслові конструкції)

- сприйняття (формування перцептивних образів на основі властивостей сприйманого об'єкта і перцептивних дій)

- розуміння (визначення сенсу, як актуальної цінності в системі значень об'єкта для суб'єкта, що має ситуативну обумовленість, яка визначається даним контекстом).

Множина предметів, яке поєднується в понятті, називається обсягом поняття (екстенсіоналом) поняття.

Щоб перейти до обсягу поняття, потрібно визначити відмінність між предметами реальними і абстрактними. Поняття, в якому мислиться ознака предмета або відношення між предметами, називається абстрактним. Сукупність абстрактних предметів і їх категорій, що відповідають одному і тому ж поняттю для яких воно є загальним, становить його обсяг, а також всі об'єкти, до яких відноситься дане поняття.

Основними методами утворення понять є:

• **Категоризація**, як когнітивний процес, – розуміється осмислення об'єктів і явищ дійсності в рамках категорій – узагальнених понять. Результатом категоризації як когнітивного процесу є формування когнітивних диференціальних ознак і когнітивних класифікаційних ознак, які виявляються як в групах концептів так і в окремих концептах. Когнітивна диференційна ознака – це окрема ознака об'єкта, усвідомлена і відображена в структурі концепту як окремий елемент його змісту. Когнітивна класифікаційна ознака – це компонент змісту концепту, що відображає той чи інший аспект, параметр категоризації відповідного об'єкта або явища і узагальнюючий однорідні диференціальні когнітивні ознаки в структурі концепту. Класифікаційні когнітивні ознаки завжди є загальними для ряду, групи або багатьох концептів. В межах концепту когнітивні класифікаційні ознаки впорядковують в єдину структуру численні диференціальні когнітивні ознаки, що утворюють зміст концепту.

• **Поділ** поняття – розбиття на декілька незалежних понять.

- **Класифікація** (ділення обсягу поняття) – розподіл об'єктів за групами на основі встановлення подібності та відмінності між ними. Різні підстави дають різні класифікації одного і того ж поняття будівництва.

- **Класифікація понять** (позиціонування класу понять будівництва до класу понять) відносить його до певної попередньо окресленої групи понять, що має певні набори ознак, характеристик і може мати окреслену поведінку, надає йому всі властивості і поведінку, які вже має даний клас понять.

- **Поняття** як семантична конструкція (СК) – відношення на ОС, включаючи специфіковані поняття, ОС – база для наповнення СК будівництва.

- **Узагальнення понять** і обмеження – логічна операція переходу від родового поняття до видового.

- **Систематизація** – процедура об'єднання, відомостей груп однорідних за деякими ознаками одиниць до певної ієрархічної єдності в функціональних цілях на основі існуючих між ними зв'язків або обраного принципу.

- При формуванні понять використовуються аналіз, розчленування змісту предмета на складові його властивості, ознаки, синтез, абстрагування, узагальнення, порівняння (порівнянню надається велика роль в рамках процесів категоризації і формування понять, зіставлення з шаблоном, з прототипом, зіставлення за ознаками відмінності, порівняння за схожими ознаками, порівнянні як структурного взаємовідображення).

- **Дефініції понять** (мається на увазі наявність властивості відношень між окремими об'єктами, які не залежать від внутрішньої структури об'єктів) – логічної операції розкриття поняття або значення терміна, за формою є фундаментальним, як кон'юнктивна-диз'юнктивна сукупність тверджень, що містять визначення і визначають поняття в цих твердженнях. В явних визначеннях задано поняття, що визначається, і, що визначає, обсяги яких є рівними і специфічними ознаками є реєструючі ознаки. Визначення понять (контекстуальне визначення, індуктивні визначення, визначення через аксіоми) є логічна операція, яка розкриває його зміст або встановлює значення терміна, яка характеризуються тим, що термін, який визначається, використовується у визначенні поняття, яке йому приписується в якості його сенсу.

- Інші способи введення понять, подібні до визначення: опис, характеристика, роз'яснення за допомогою прикладу.

Поняття нерозривно пов'язане з основною **мовною одиницею** – словом або словами і словосполученнями. Утворена сукупність аксіоматичних смислових одиниць тексту визначає поняття, що погоджує сукупність інформаційних одиниць з перерахованими вище операціями, зі здоровим глуздом і значенням. До основних результатів контент-аналізу відноситься:

- класифікація смислових одиниць прикладної галузі будівництва в рамках якої вони визначаються;
- кластеризація смислових одиниць, яка здійснюється відповідно до їх приналежності до галузі будівництва в рамках якої вона використовується;
- багаторівнева структуризація тексту, що забезпечує вилучення смислів як завдання пошуку пов'язаних ознак;
- проспекція текстів, яка проявляється в передбаченні змістовно-фактуальної інформації, яка експліцитно виражена в наступних частинах тексту і представлена хронологічно;
- інтегративність текстів модуля знань, що проявляється в семантичному і формально-структурному аспектах функціонування тексту як єдиного цілого.

Текст знань утворюється за рахунок текстуальних зв'язків, текстоутворюючих засобів, словосполучень, контексту і характеризується синтаксичними, семантичними та лінгвістичними параметрами. Зміст знань – це інформація тексту знань і відповідно до неї здійснюється класифікація за типами інформації тексту знань:

- змістовно-фактуальною;
- змістовно-концептуальною;
- змістовно-підтекстовою.

Мінімальною структурною одиницею знань є інформаційна одиниця – сукупність виражених мовними засобами смислових одиниць.

Термін – це слово або словосполучення, що позначає строго визначене поняття і характеризується однозначністю в межах спорідненої групи наук [23,24]. Терміни існують в рамках певної термінології, тобто входять в конкретну лексичну систему мови за посередництвом конкретної термінологічної системи будівельної галузі, забезпечуючи наочність судження. Таким чином, основною специфічною рисою онтології будівництва є широке використання термінології. У той же час, термінологія не ізольована від природної мови і в ній присутня багатозначність терміну через різні відтінки значення або авторського переосмислення вже існуючих термінів або між науковою термінологічною омонімією [12,13,14].

В рамках тексту знань має місце наступний розподіл лінгвістичних понять за структурними одиницями:

- «зміст» (*Cnt*) – це морфеми, слова, словосполучення, синтаксичні конструкції знань. Вимірність поняття «зміст» заснована на інформативності елементів тексту (слів, пропозицій, фраз) і частоті їх появи;
- «сенс» (*Sns*) – це думка, повідомлення, що є в реченні, які містяться в тексті в цілому. Вимірність поняття «сенс» засноване на інформативності та позиціонуванні термінів в тексті (*Aims*);

- «значення» (Vlu) – сукупність знань для яких об'єднаний текст характеризується відносною смисловою завершеністю. Вимірність поняття «значення» базується на повноті присутності в тексті спеціальних слів і словосполучень – маркерів важливості, які характеризують їх змістовну значимість.

Нехай $D(Cnt) = D(Sns) = D(Vlu)$ – синтаксичні конструкції як область визначення D відповідних функцій і $B(D)$ – бульян від D .

Тоді, відповідно: $Cnt: B(D) [0,1] = E(Cnt)$ область значень Cnt ;

$Sns: B(D) Aims = E(Sns)$ область значень Sns ;

$Vlu: B(D) B(D) = E(Vlu)$ область значень Vlu ,

$\{F\}: \{f: B^3(D) E(Cnt) E(Sns) E(Vlu)\}$ породжують семантичні структури на класі знань.

Семантична структура, що розглядається як ієрархія смислів, є організованою замкненою системою, яка при зміні ряду своїх ознак і під впливом зовнішніх чинників, змінює конфігурацію і стає новою замкненою системою.

Новацією запропонованої і експериментально перевіреної моделі є метод виділення з тексту змістовно-значущих одиниць і перехід до моделей.

Операційний аналіз поняття «введення знань» в семантичній структуризації модуля знань є актуальним завданням, що обумовлено необхідністю уніфікації уявлень знань.

Завдання операційного аналізу поняття в семантичній структурі знань – це побудова представлення структури перетворень семантичних значень в рамках засобів граматичних форм подачі всіх його реалізованих семантичних конфігурацій.

Під поняттям розуміємо текст, що складається з декількох пов'язаних між собою речень-висловлювань, що має певну змістову, щодо деякого обсягу фактів, інформацію або інформаційні одиниці. В рамках тексту знань має місце наступний розподіл лінгвістичних понять за структурними одиницями: $\{Cnt, Sns, Vlu\}$, що породжують семантичні структури даних знань.

Нехай L – кількість понять які вводяться у модулі – зв'язаної множини процедур визначення понять, l – оператор, відповідний введенню поняття, і вся множина L впорядкована і пронумерована, тобто:

- впорядкування сукупності смислових одиниць (i) при введенні l -ю процедурою поняття у k -му модулі: $\forall i, j \forall l, k, i < j : CE_i^{lk} \prec CE_j^{lk}$;
- впорядкування процедур введення l_1, l_2 по відношенню до модуля з номером k : $\forall i, j, k \forall l_1 < l_2 : CE_i^{l_1 k} \prec CE_j^{l_2 k}$;
- впорядкування між собою смислових одиниць модулів з номерами k_1, k_2 : $\forall i, j \forall l_1, l_2 \forall k_1 < k_2 : CE_i^{l_1 k_1} \prec CE_j^{l_2 k_2}$.

Тоді:

- $|L| < \infty$;
- $\forall l \in L \exists l_i \in N : l \leftrightarrow l_i$ – вводяться одночасно;
- $L \equiv \{l\}$ – частково впорядкована множина.

Нехай CE_i^l – i -та смислова одиниця l , IE_i^l – i -та текстова одиниця l ,
тоді:

- $\{CE_i^l\} = \{CE_{i_j}^l\} \cup \{CE_{i_k}^l\}$ – сукупність смислових одиниць l ;
- $\{CE_{i_j}^l\}$ – сукупність смислових одиниць l , які визначені раніше;
- $\{CE_{i_k}^l\}$ – сукупність смислових одиниць l , що первинно визначені,
 $\forall l, \forall i, \forall (j \neq k) : CE_{i_j}^l \neq CE_{i_k}^l$;
- $\forall i \forall l : CE_i^l \leftrightarrow IE_i^l$ – ізоморфізм для точних наук.
- $\{CE_i^{*l}\}$ – сукупність смислових одиниць l , що визначаються аксіоматично:

$$\forall i \forall j \forall l \forall m (l \neq m) : CE_i^{*l} \neq CE_j^{*m}$$

$$\forall l : \{CE_i\}_l \rightarrow \{CE_i\}^l;$$

$\{CE_i\}_l$ – сукупність вихідних смислових одиниць l ;

$\{CE_i\}^l$ – сукупність всіх смислових одиниць l і виконуються умови:

$$\{CE_i\}_l \subseteq \{CE_i\}^l; |\{CE_i\}_l| \leq |\{CE_i\}^l| < \infty;$$

$\forall l, \{\{CE_i\}_l, \{CE_i\}^l\}$ – частково впорядковані;

$\{\{CE_i\}_l \cup_{l \in L} \{CE_i\}^l\}$ – частково впорядковані;

$$\left| \{ \{CE_i\}_l \cup_{l \in L} \{CE_i\}^l \} \right| < \infty; \forall l, \forall m \geq 1 : \{CE_{i_k}^{l+m}\} \cap \{CE_{i_k}^l\} = \emptyset.$$

Тоді, ефективність $Ef(\bullet)$ обраних аксіоматичних одиниць $Ef\left(\bigcup_{l \in L} CE_i^{*l}\right)$ визначається як:

$$Ef\left(\bigcup_{l \in L} CE_i^{*l}\right) = \frac{\left| \bigcup_{l \in L} CE_i^l \setminus \bigcup_{l \in L} CE_i^{*l} \right|}{\left| \bigcup_{l \in L} CE_i^{*l} \right|}.$$

Нехай відповідно до визначення поняття (смислової одиниці) $\{P_r\}_{i=1}^n$ – сукупність основних ознак, що утворюють поняття (смыслові одиниці) $\bigcup_{l \in L} CE_i^{*l}$.

Оцінка обсягу V різних між собою понять щодо знань предметної області: $V = wdt(l) \times dpt(l)$, де $wdt(l) = |\{CE_{i_k}^l\}| / 2^n$; $dpt(l) = \{\text{середня довжина ланцюжка введення } l\text{-го поняття}\}$.

Тим самим: операційний аналіз поняття в семантичній структуризації знань показав, що можливо ввести заходи щодо уніфікації уявлень знань, які дозволяють будувати кількісні оцінки.

Визначимо основні особливості понять в будівництві:

- a) частина понять відноситься до реальних, частина – до абстрактних об'єктів;
- b) частина понять відноситься до понять СБ, а частина – до будівництва в цілому;
- c) поняття є основою для комунікації всередині та поза для СБ;
- d) поняття є засіб подання інформаційного об'єкта будівництва;
- e) поняття у різних СБ утворюються не синхронно;
- f) сукупність близьких понять у різних СБ – основа формування поняття будівництва;
- g) виразність понять у різних СБ різна;
- h) охоплення сутності явища за рахунок понять СБ в різних СБ різні;
- і) наявність і обумовленість еволюції у понять будівництва.

Сукупність понять в ОС змінюється за рахунок:

- зміни достовірності і актуальності понять в ІСПОБ;
- інтенсивності і масштабності діалогу і поля завдань в ІСПОБ;
- вилучення поняття, ключових слів, пов'язаних слів, рефератів, шаблонів з текстів;
- розширення представлення поняття семантичними концептуальними графами (КГ);
- розширення представлення поняття онтологічними графами (ОГ);
- багатокритеріального методу експертного оцінювання поняття в умовах нечіткої інформації, можливістю поповнити інформацію про предмет експертизи;
- процесами утворення архаїзмів в ОС.

Лінгво-семантичний аналізатор – основний метод аналізу появи нових понять: ототожнення семантично сумісних понять з різних джерел знань СБ, формування текстової стереотипності, прототипів і інваріантів у вигляді тексту, що з точки зору тезауруса і словника написаний словами даного словника і термінами, які містяться в тезаурусі, – за рахунок пошуку в текстових масивах різного ступеня структуризації мінімальних смислових одиниць і використанні об'єктно-орієнтованого багатоваріантного аналізу в побудові системи шаблонів з недовизначеними обчисленнями для полів класів, де заходи семантичної близькості припускають однозначну інтерпретацію термів для однієї онтології, порівняння таксономій понять і інших відношень різних онтологій.

Процедура введення або отримання поняття визначається наявністю інформації щодо наступного:

- 1) простору ознак процедури оцінки знань:
 - 1.1 семіоми (терміни, словосполучення, ідіоми);

1.2 інформаційні одиниці

- 2) подання ознак простору процедури оцінки знань;
- 3) денотатних графів процедури оцінки знань;
- 4) концептуальних графів процедури оцінки знань;
- 5) семантичного графу тексту процедури оцінки знань;
- 6) семантичної мережі як уявлення тексту процедури оцінки знань;
- 7) специфіки уявлення ознак простору процедури оцінки знань графами:
 - 7.1 словник і тезаурус;
 - 7.2 словник процедури оцінки достовірності знань;
 - 7.3 словник фразеологізмів;
- 8) оцінка як міра на графі (семантичної мережі):
 - 8.1 схожість графів;
 - 8.2 навантаженість ребер графа як міри важливості;
 - 8.3 не лінійність оцінки як міри;
 - 8.4 залежність шкали оцінки від складності, інформативності, інтенсивності появи нових понять;
 - 8.5 області визначення оцінки в залежності від обсягу отриманих текстів і зростання словника фразеологізмів;
 - 8.6 динаміки смислових одиниць в залежності від обсягу отриманих текстів нових понять;
 - 8.7 мовний поріг знань і умінь СБ;
 - 8.8 уміння і навичок викладення при формалізації СБ;
 - 8.9 межі компетентності СБ.

Поняття у формальній мові вводиться як «Це є (*def*)» або «Це є ознаки (\subseteq)» або за допомогою операції конкатенації.

На множині ознак існують відношення: ознаки \rightarrow поняття – $fix\ i \Rightarrow P_i \Rightarrow I_{P_i}$.

Множина об'єктів носіїв ознаки P_i визначається як:

$$\{\exists\{\forall x \in I_{P_i} | x \leftrightarrow n_x \in N\} \exists N_i \geq n_x\} \wedge \{|Ob| \leq N_0(\infty)\}.$$

Тоді $\{Ob\} = \text{support}\varphi$, де $\varphi: \{Ob\} \equiv \{\varphi^{-1}(I_{P_i} = x) | \{x\} \equiv \text{automorf } I_{P_i}\}$.

Об'єкти з індексами k, l роздільні за ознакою P_i , якщо існує непустий переріз шкал ознак $IP_i \neq 0$ де $IP_i^k \cap IP_i^l = \emptyset$.

Об'єкти ob_l, ob_k неможна порівняти за ознакою P_i , якщо $\{\exists P_i(ob_l) \wedge \neg \exists P_i(ob_k)\}$.

Структурою на $\{Ob\}$ є розподіл ознак за екземплярами об'єктів: $S_{\{P_i\}}(Ob) = \times_i S_{P_i}(Ob)$.

Потужність множини структур становить:

$$|S_{\{P_i\}}(Ob)| = \prod_{i=1}^n 2^{N_0} N_i! = 2^{nN_0} \prod_{i=1}^n N_i!.$$

Нехай P_i – поняття i -ї процедури його введення, P_i^* – базові поняття в цій процедурі, P_i^j , ($j < i$) – поняття з j -ї процедури його введення, які використовуються в i -й процедурі. Тоді, $P_i = P_i^0 \cup P_i^* \bigcup_{j=1}^{i-1} P_i^j$, причому $\{P_i\}$ – частково впорядкована множина, де:

1. $P_i \subseteq P_i$ – рефлексивність;
2. $\frac{P_n \subseteq P_i}{P_i \subseteq P_n} \Rightarrow P_i \equiv P_n$ – анти симетричність;
3. $P_n \subseteq P_i \subseteq P_k \Rightarrow P_n \subseteq P_k$ – транзитивність;
4. $\forall i > n: \left\{ \left\{ P_n \subseteq P_i \Rightarrow P_n^0 \cup P_n^k \bigcup_{k=1}^{n-1} P_i^k \subseteq P_i^0 \cup P_i^* \bigcup_{k=1}^{i-1} P_i^k \Rightarrow \exists \tilde{P}_i \notin P_n \right\} \wedge \left\{ P_i \subseteq P_n \right\} \right\}$.

З точністю до формальних денотатів, існує п'ять схем реалізації процедури введення поняття. Позначимо оператор введення понять як D , i, k – номери ланцюгів введення понять, перший нижній індекс – коренева точка, другий нижній індекс – номер кроку по ланцюгу.

D_{j0}^i – аксіоматичні визначення;

D_{0j}^i – виведення;

D_0^i – кореневі аксіоматичні визначення;

$\bigcup_i D_0^i[0, n_i]$ – поняття процедури введення;

$D_0^{i*} \equiv \left\{ \bigcup_i D_0^i[0, n_i] \right\}$ – вихідні поняття процедури введення;

$D_{0np}^{i*} \equiv \left\{ \bigcup_i D_0^i[0, n_i] \setminus D_0^{i*} \right\}$ – проміжні поняття процедури введення.

D_{0j}^{i*} , D_{j0}^{i*} , D_0^{i*} – введення або отримання понять, які в подальшому не використовуються.

У введених позначеннях справедливі наступні твердження:

- 1) $\forall i \exists D_0^{i*}$ – існують ланцюги виводу;
- 2) для будь-якого виводу* – існує процедура виведення від кореня $D_0^k[0, j]$;
- 3) дерево виводу $G = \bigcup_i G(D_0^i)$, де $G(D_0^i)$ – i -те дерево виводу від кореня;
- 4) $\forall i, j D_0^{i*} \cap D_0^{j*} = \emptyset$;
- 5) G – зв'язний граф, який не містить автономних процедур виведення.

Тоді, п'ять схем реалізації процедури введення поняття можна представити у вигляді рис 14-17.

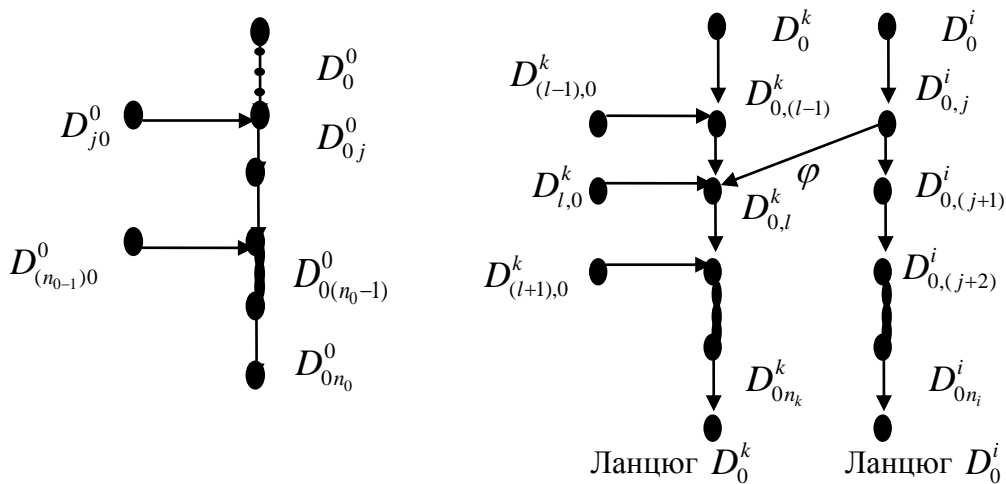


Рис. 14 Схеми реалізації процедури введення поняття, де:
 $dt(D_{0,j}^i) = dt(D_{j,0}^i) + dt(D_{j-1,j}^i)$, $dt(D_{0,l}^k) = dt(D_{l,0}^k) + dt(D_{l-1,l}^k + dt(\varphi(D_{0,j-1}^i)))$

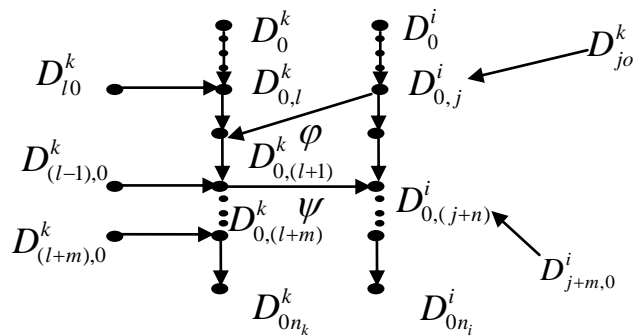


Рис. 15 Схеми реалізації процедури введення поняття, де:
 $dt(D_{0,j+n}^i) = dt(D_{j,i+n}^i) + dt(D_{j+m,0}^i) + dt(\psi(D_{0,l+m}^k))$
 $\{dt(D_{0,l+1}^k) + dt(D_{l+1,0}^k) + dt(\varphi(D_{0,j}^i)) + dt(D_{0(l+1,l+m)}^k) + dt(D_{l+m,0}^k)\}$

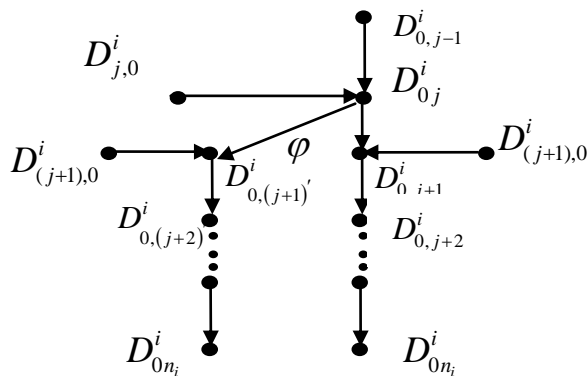


Рис. 16 Схеми реалізації процедури введення поняття, де:
 $dt(D_{0,j'}^i) = dt(D_{j',0}^i) + dt(\varphi(D_{0,j+1}^i))$, $dt(D_{0,j}^i) = dt(D_{j+1,0}^i) + dt(D_{j-1,l+1}^i)$,
 символ «'» – нумерація вздовж нового підланцюга

На схемі, представлений на рис. 16 процедура φ виконує наступні операції:

- 1) розщеплення системи понять – подачу процедури введення:
 $D[\varphi(D_{0,j-1}^i)] \cap D[D_{j-1,j}^i] = \emptyset$;
- 2) розгалуження системи понять і подачу процедури введення:
 $D[\varphi(D_{0,j-1}^i)] \cap D[D_{j-1,j}^i] \neq \emptyset$;
- 3) розбіжність подачі процедури введення: $D[\varphi(D_{0,j-1}^i)] = D[D_{j-1,j}^i]$.

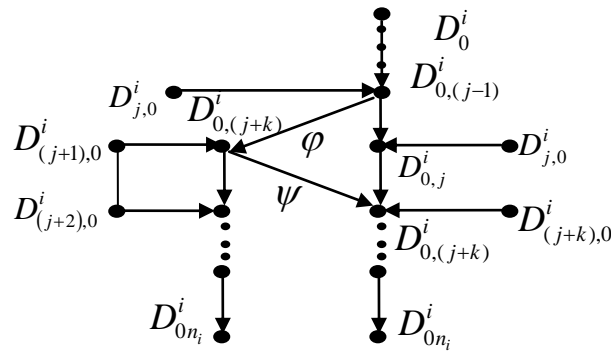


Рис. 17 Схема реалізації процедури введення поняття де для $D_{0,j}^i$, $D_{j,0}^i$, D_0^i виконуються наступні співвідношення: $|dt(D_{0,j}^k)| \leq |dt(D_{0,j-1}^k)| + |dt(D_{0,l}^k)|$

На схемі представлений на рис. 17 для $D_{0,j}^i$, $D_{j,0}^i$, D_0^i виконуються наступні співвідношення:

- 1) $|dt(D_{0,j}^k)| \leq |dt(D_{0,j-1}^k)| + |dt(D_{0,l}^k)|$;
- 2) $D_{0,l}^i \setminus \varphi^{-1}(D_{0,j}^k) = \Delta\varphi_{l,j}^{i,k}$;
- 3) $dt(D_{0,j+k}^i) = dt(D_{j+k,0}^i) + dt(\psi(D_{0,(j+k)}^i) + D(D_{j-1,j+k}^i)) =$
 $= dt(D_{j+k,0}^i) + D(D_{j-1,j+k}^i) + dt(\psi(D_{j',0}^i) + dt(\varphi(D_{0,j-1}^i)))$.

Для понять оголошених денотатів D_0^i і денотатів, що формально приводяться $D_{j,0}^i$, з метою їх подальшого оголошення в $D_{0,j}^i$ нижченаведена схема (рис. 18) ілюструє їх введення:

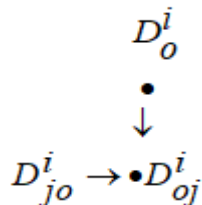


Рис. 18 Схема введення понять

$$\text{Нехай: } D_{0,j}^i \begin{array}{c} \xrightarrow{\varphi} \\ \xleftarrow{\psi} \end{array} D_{0,l}^k, \quad i \neq k.$$

Тоді справедливі наступні співвідношення:

$$\begin{aligned} & 1) \\ & \forall i, k, j, l (i \neq k): \neg \exists \varphi, \psi : \{ \varphi(D_{0,l}^k) \subseteq D_{0,j}^i \} \cup \{ \psi(D_{0,j}^i) \subseteq D_{0,l}^k \} \Rightarrow \\ & \Rightarrow \{ D_0^i \cap D_0^k = \emptyset \} \cup \{ D_0^i \equiv D_0^k, B(D_0^i) = B(D_0^k) \cup B(D_0^{i_2}), B(D_0^i) \cap B(D_0^{i_2}) \equiv \emptyset \} \\ & \quad \neg \exists j, l (i \neq k): \neg \exists \psi, \exists \varphi : dt(D_{0,j}^i) \supseteq dt(\varphi(D_{0,j}^i)) \Rightarrow \\ & 2) \Rightarrow \{ \{ dt(D_{0,j-1}^i) \cap dt(D_{0,l-1}^k) = \emptyset \} \wedge \{ D_0^i \cap D_0^k = \emptyset \} \} \cup \\ & \quad \cup \left\{ \{ D_0^i \cap D_0^k \neq \emptyset \} \wedge \left\{ \begin{array}{l} B(D_0^i \setminus D_0^k) \supset dt(D_{0,j-1}^i) \wedge \\ \wedge B(D_0^k \setminus D_0^i) \supset dt(D_{0,l-1}^k) \end{array} \right\} \right\} \\ & \quad \neg \exists j, l (i \neq k): \exists \psi, \varphi : dt(\psi(D_{0,j}^i)) \subseteq dt(D_{0,l}^k) \Rightarrow \\ & 3) \Rightarrow \{ \{ D_0^i \cap D_0^k \neq \emptyset \} \wedge \{ dt(D_{0,j-1}^i) \cap dt(D_{0,l-1}^k) = \emptyset \} \} \cup \\ & \quad \cup \{ \{ D_0^i \cap D_0^k \neq \emptyset \} \wedge B(D_0^i \setminus D_0^k) \supset dt(D_{0,j-1}^i) \wedge B(D_0^k \setminus D_0^i) \supset dt(D_{0,l-1}^k) \} \} \end{aligned}$$

Таким чином для ОС будівництва:

- сприйняття – це функція, яка полягає в перетворенні фрагмента реальності в ситуаційний предикат,
- розуміння – це функція, яка полягає в перетворенні тексту в смисловий предикат,
- мислення – це функція, яка полягає в обчисленні значення смислового предиката на підставі його аргументу – ситуаційного предиката.

Вхідним аргументом функції мислення є:

- суб'єктивна модель зовнішнього світу – ситуація і смисловий предикат – зміст тексту,
- результатом – значення смислового предиката в даній ситуації;
- синтаксис пропозиції – структура формули предиката,
- сенс цієї пропозиції - функція реалізована предикатом.

При організації досліджень в такій предметній області як будівництво, з інтенсивним використанням даних доцільно визначати специфікації предметної області, що включає визначення понять предметної області засобами онтологій, так як вони не є системами гільбертового типу (коли докази виходять формальним чином з аксіом шляхом застосування правил виведення), що не є структурованими. Концептуальні граfi є семантичними моделями тексту предметних областей з інтенсивним використанням даних, і моделюють структури, наділені змістом: під моделлю розуміється дводольний денотатний граф, на якому в свою черга, прописаний концептуальний граф.

Ефективне застосування концептуальних графів пов'язано з вирішенням завдань агрегування і кластеризації на графах. Математична модель смислового предиката у вигляді смислового графа із спрямованими зв'язками реалізується на основі семантичної нейронної мережі, як способу обробки і аналізу даних будівництва [9,10, 25]. У цьому випадку операції розпізнання значень предметних змінних моделюються в неоднорідній нечіткій семантичній нейронній мережі. Семантична нейронна мережа дозволяє обробляти зміст тексту як функцію деякої алгебри і смисловий граф як математична абстракція може бути повністю нею змодельований. Окремий нейрон позначає елементарне поняття аналізованої мови, послідовність застосування операцій задається структурою зв'язків між нейронами, градієнтні величини – відношення між поняттями при нечіткій та неповній інформації. При цьому сенс – миттєвий стан частини нейронної мережі, відповідальної за вилучення інформації. Модель процесу формування поняття – це нейромережа (НМ) як механізм подання його когнітиву в предметній області будівництва [28,29,30]. Моделювання врахування досвіду і нових знань для створення технологій власне в будівництві на основі нейромереж над сукупністю кортежів: {(проект, локація, суб'єкт будівництва, МТР, зовнішнє середовище, режим експлуатації, етап життєвого циклу)}. До переваг і недоліків НС для моделювання формування поняття відноситься:

- переваги:
 - автоматичне набуття нових знань;
 - навчання;
 - виконання узагальнень;
- недоліки:
 - перенавчання. Воно полягає в тому, що мережа «запам'ятовує» відповіді замість того, щоб вловлювати закономірності в даних;
 - необхідність нормалізації батчів;
 - вимоги до репрезентативності і несуперечливості набору даних для навчання;
- супутні проблеми:
 - підбір параметрів для нарощування вибірки, критичне збільшення часу навчання та інші;
 - пошук аномалій в навчальній вибірці;
 - катастрофічна «забудькуватість»;
 - доповнення до НМ для моделювання формування поняття в будівництві;
 - метанейромережа (МНМ) над сукупністю НМ також породжує поняття;
 - специфіка і завдання МНМ – навчання НМ як об'єкту;
 - висновок (мислення) у МНМ це вибір та класифікатори.

Висновки. На основі вище сказаного, визначена процедура моделювання співвідношення когнітивних і семантичних процесів в будівництві та побудова моделі утворення нових понять:

1. Виділення об'єкта, що спостерігається декларативно, на основі ознак або заходів, і тим самим, визначення його інформаційного простору.
2. Виділяємо «схожі об'єкти» в інформаційному просторі з урахуванням їх прив'язки до процесів будівництва.
3. Визначаємо семантичне ядро для безлічі «схожих об'єктів».
4. Побудова передбачуваний клас еквівалентності на основі семантичного ядра і для безлічі «схожих об'єктів».
5. Визначення прототипу з числа «схожих об'єктів» для побудови класу об'єктів.
6. Визначення нового поняття, його обсягу та змісту на класі об'єктів, визначених прототипом.
7. Побудована схема діалогу ІСПОБ і схема сполученої діалогової завдання
8. Доведено твердження про існування рішення ОДЗ ІСПОБ і спряженої ОДЗ;
9. Досліджено процес діалогової трансформації ІСПОБ за рахунок формулювання і перевірки гіпотез.

Література

1. Большакова Е.И. Автоматическая обработка текстов на естественном языке и анализ данных: учеб.пособие / Е.И. Большакова, К.В. Воронцов, Н.Э. Ефремова, Э.С. Клышинский, Н.В. Лукашевич, А. С. Сапин. Москва : Изд-во НИУВШЭ, 2017. 269 с.
2. Горда О.В. Застосування ВІМ технологій на будівельному майданчику в інформаційних технологіях управління проектом. Міжнародна науково-практична конференція молодих вчених «БУД-МАЙСТЕР-КЛАС-2019», 29-31 листопада. Київ :КНУБА, 2019. С. 424-425.
3. Григоровський П. Є., Горда О. В., Чуканова Н. П. Інформаційні середовища в будівництві. // *Будівельне виробництво*. № 68. 2019. С. 15-19.
4. Горда О.В. Аналіз моделей в інформаційному просторі будівництва. Міжнародна науково-практична конференція молодих вчених «БУД-МАЙСТЕР-КЛАС-2020», 25-27 листопада 2020. КНУБА. С. 306-308.
5. Горда О.В. Специфіка інформаційних середовищ в будівництві. VII міжнародна науково-практична конференція «Управління розвитком технологій». Київ : КНУБА, 2020. С. 55-56.
6. Горда О.В. Топологія інформаційного простору в будівництві / *Будівельне виробництво*. Київ, 2020. Вип. 70. С. 39-44.
7. Горда О.В. Поле задач об'єкту будівництва / *Управління розвитком складних систем*. Київ : КНУБА, 2020. Вип. 44. С. 78-83.

8. Чехарин Е.Е. Информационная модель семантического окружения / *Перспективы науки и образования*. 2014. Вип. 4. С. 20-24.
9. Хайкин С. Нейронные сети: полный курс. Москва : Вильямс, 2006. 1104 с.
10. Рашид Т. Создаем нейронную сеть. Москва : Диалектика, 2019. 272 с.
11. Андрич О.Ф., Макушкина Л.А. Исследование методов оценки качества готовых онтологических моделей / *Современные научные исследования и инновации*. 2014. Вип. 3 (35). С. 11.
12. Береснев С.Д., Фалькова В.Ю. Соотношение признаков стандартности и авторской индивидуальности научно-технического текста. Разновидности текста в функционально-стилевом аспекте. Пермь : Перм. ун-т, 1994. С. 41–51.
13. Котюрова М.П. Многоаспектность явлений стереотипности в научных текстах, текст: стереотип и творчество. Пермь : Перм. ун-т, 1998. С. 5-30.
14. Харин Ю.А., Зеленков А.И., Творчество: критика и преемственность, Творчество в научном познании. // Под общ. ред. Д.И. Широконова, Ю.А. Харина. Минск : Наука и техника, 1976. 232с. .
15. Никитина С.Е. Семантический анализ языка науки. На материале лингвистики. Москва : Наука. 1987. 126 с..
16. Осипов Г.С. Приобретение знаний интеллектуальными системами Основы теории и технологии. Москва : Наука, 1997. 112 с.
17. Моделирование языковой деятельности в интеллектуальных системах. Под ред. А.Е. Кибрика и А.С. Нариньяни. Москва : Наука, 1987. 280 с.
18. Соломатин Н.М. Информационные семантические системы // В уч. пос. Перспективы развития вычислительной техники в 11 кн. Кн. 1. Москва : Высшая школа, 1989. 130 с.
19. Колшанский Г.В. Контекстная семантика. Москва : Наука, 1980. 154 с.
20. Андрич О.Ф., Макушкина Л.А. Исследование методов оценки качества готовых онтологических моделей / *Современные научные исследования и инновации*. 2014. Вып. 3 (35). С. 11.
21. Гинзбург С. Математическая теория контекстно-свободных языков. Москва : Из-во "Мир", 1970. 326 с.
22. Строительство: Энциклопедический словарь / Автор-составитель Артюхович Д.В. Ставрополь: Ставропольское издательство «Параграф», 2011. 766 с.
23. Алексеева Л. М. Проблемы термина и терминообразования. Пермь : Перм. ун-т, 1998. 120 с.
24. Мельникова Е. Н. Сигнификативное значение предложения при рассмотрении его семантико-синтаксической структуры / *Лингвистика XXI века: традиции и новации*. 2016. С. 145-150.
25. Дударь З.В., Шуклин Д.Е. Семантическая нейронная сеть, как формальный язык описания и обработки смысла текстов на естественном языке / *Радиоэлектроника и информатика*. 2000. Вып. 3. С. 72-76.
26. Шабанов-Кушнаренко Ю.П. Теория интеллекта. Математические средства. Харьков: Вища школа, 1984. 143 с.

27. Москальчук Г. Г. Структура текста как синергетический процесс
Издательство: "Едиториал УРСС", 2010. 296 с.
28. Бочаров В. А., Маркин В. И. Введение в логику: Учебник. Москва :
ИД «ФОРУМ»: ИНФРА-М, 2010. 560 с.
29. Бенерджи Р. Теория решения задач. Подход к созданию
искусственного интеллекта. Москва : Мир, 1972. 189 с.
30. Носков А.А. Метод выделения в тексте конструкций по их лексико-
синтаксическим шаблонам // Сборник статей молодых ученых факультета
ВМиК МГУ. Москва : Издательский отдел фак-та ВМиК МГУ им. М.В.
Ломоносова; МАКС Пресс, 2009. Вып. 6. С. 136-145.

References

1. Bolshakova E.I. (2017.) Automatic processing of texts in natural language
and data analysis: textbook / E.I. Bolshakova, K.V. Vorontsov, N.E. Efremova,
E.S. Klyshinsky, N.V. Lukashevich, A.S. Sapin. Moscow : Publishing House of
Higher School of Economics, 269 p.
2. Gorda O.V. (2019) Stagnation of BIM technologies on the alarm maidan in
the information technologies of project management. / International scientific
and practical conference of young students "BUD-MAISTER-CLAS-2019", 29-
31 leaves. Kuiv : KNUBA. P. 424-425.
3. Grigorovskiy P.Y., Gorda O.V., Chukanova N.P. (2019) Building information
environment. / Budivelne virobnitstvo. No. 68. P. 15-19.
4. Gorda O.V. (2020) Analysis of models in the information space of business. /
International scientific and practical conference of young scientists "BUD-
MAISTER-CLAS-2020". Kuiv : KNUBA. P. 306-308.
5. Gorda O.V. (2020) The specifics of the information environments in
construction. / VII international scientific-practical conference "Management of
development of technologies". Kiev : KNUBA. P. 55-56.
6. Gorda O.V. (2020) Topology of information space in construction. /
Budivelnaya virobnitstvo, №70. P. 39-44.
7. Gorda O.V. (2020) The task field of the construction site. / Management of
complex systems development. No 44. P. 78-83.
8. Chekharin E.E. (. 2014) Information model of the semantic environment /
Prospects for Science and Education. No. 4. P.20-24.
9. Haykin S. (2006) Neural networks: a complete course. Moscow : Williams.1104 p.
10. Rashid T. (2019) Create a neural network. Moscow : Dialectics. 272 p.
11. Andrich O.F., Makushkina L.A. (2014) Research of methods of estimation
of quality of ready ontological models / Modern scientific researches and
innovations. No 3 (35). P. 11.
12. Beresnev S.D., Falkova V.Yu. (1994) The ratio of features of standard and
authorial individuality of scientific and technical text. Varieties of text in the
functional and stylistic aspect. Perm : Perm. Univ. P. 41–51.

13. Kotyurova M.P. (1998) Multifaceted phenomena of stereotypes in scientific texts, text: stereotype and creativity. Perm: Perm. Univ. P. 5–30.
14. Kharin Y.A., Zelenkov A.I. (1976) Creativity: criticism and continuity, Creativity in scientific knowledge. / Under common. ed. D.I. Shirokanova, Yu.A. Harina. Minsk : Nauka i tehnika,. 232 p. .
15. Nikitina S.E. (1987) Semantic analysis of the language of science. On the material of linguistics. Moscow : Science. 126 p.
16. Osipov G.S. (1997) Acquisition of knowledge by intelligent systems Fundamentals of theory and technology. Moscow : Nauka. 112 p.
17. Modeling of language activity in intelligent systems. Ed. A.E. Kibrika and A.S. Narinyani. Moscow : Nauka. 280 p.
18. Solomatin N.M. (1989) Information semantic systems. - // In the account. pos. Prospects for the development of computer technology in 11 books. Book 1. Moscow: Higher school. 130 p.
19. Kolshansky G.V. (1980) Contextual semantics. Moscow : Nauka. 154 p.
20. Andrich O.F., Makushkina L.A. (2014) Research of methods of estimation of quality of ready ontological models / Modern scientific researches and innovations. No 3 (35). P. 11.
21. Ginzburg S. (1970) Mathematical theory of context-free languages. Moscow : Publishing house "Mir". 326 p.
22. Construction: Encyclopedic Dictionary. (2011) / Author-compiler Artyukhovich D.V. Stavropol: Stavropol Publishing House "Paragraph". 766 p.
23. Alekseeva L.M. (1998) Problems of term and term formation. Perm : Perm. University. 120 p.
24. Melnikova E.N. (2016) Significant meaning of the sentence when considering its semantic-syntactic structure / Linguistics of the XXI century: traditions and innovations. P. 145-150.
25. Dudar Z.V., Shuklin D.E. (2000) Semantic neural network as a formal language for describing and processing the meaning of texts in natural language / Radioelectronics and Informatics. No 3. P. 72-76.
26. Shabanov-Kushnarenko Yu.P. (1984) Intelligence theory. Mathematical tools. Kharkov : Higher school,. 143 p.
27. Moskalchuk G.G. (2010) Text structure as a synergetic process. Publisher: "Editorial URSS". 296 p.
28. Bocharov V.A., Markin V.I. (2010) Introduction to logic: Textbook. Moscow : PH «FORUM»: INFRA-M. 560 p.
29. Benerjee R. (1972) Theory of problem solving. An approach to creating artificial intelligence. Moscow : Publishing house "Mir". 189 p.
30. Noskov A.A. (2009) Method of selection in the text of constructions according to their lexical-syntactic templates / Collection of articles of young scientists of the faculty of VMiK MSU. Moscow : Publishing department of the faculty of VMiK MSU. M.V. Lomonosov; MAX Press. Issue 6. P. 136-145.

АНАЛИЗ ЗАДАЧИ ПОСТРОЕНИЯ ОНТОЛОГИЧЕСКОГО СЛОВАРА СТРОИТЕЛЬСТВА

Рассматривается проблема автоматизированного построения онтологий в строительстве на основе семантического анализа текстов на естественном языке. Важнейшей проблемой развития исследований в строительстве является его интеллектуализация, и связанные с этим интеграция данных и качественный поиск информации. Однако данные технологии предполагают наличие качественных источников семантических данных. В целом можно утверждать, что повторное использование и интеграция с источниками данных находится на низком уровне. Это связано с тем, что источники данных не достаточно совершенны для использования или интеграции. В связи с этой проблемой, задача автоматизированного формирования онтологий на основе анализа текстов на естественном языке в строительстве достаточно актуальна. Распространенный критерий качества онтологии основан на оценке удобства и эффективности работы и связи когнитивных процессов с семантическими, с целью ее построения и актуализации в сфере строительства. Оценка построенных онтологий на основе автоматизированного инструментария является еще одной отдельной сложной задачей. Основным отличием проведенного исследования, результаты которого приводятся в данной работе, является когнитивно-семантический анализ на основе теории категорий, математической логики и универсальной алгебры, алгебры множеств, алгебры кортежей и реляционной алгебры, а именно – построение онтологического словаря строительства, онтологических конструкций на открытых языках, представление знаний в строительстве. В работе произведена классификация и определение информационных единиц и их источников. Определены особенности информационного ресурса в строительной отрасли.

Ключевые слова: строительство; онтология; понятие; тезаурус; информационный объект; изоморфизм; анализ; семантическая модель.

Ph.D., Associate Professor **Gorda E.V.**
anaelg@ukr.net, ORCID: 0000-0001-7380-0533
Kyiv National University of Construction and Architecture» (KNUCA)

ANALYSIS OF THE PROBLEM OF BUILDING ONTOLOGICAL DICTIONARY OF CONSTRUCTION

The problem of the automated construction of ontology in construction based on the semantic analysis of texts in natural language is considered. The most important problem in the development of research in construction is its intellectualization, and the associated data integration and high-quality information retrieval. However, these technologies presuppose the presence of high-quality sources of semantic data. Overall, it can be argued that reuse and integration with data sources is low. This is because the data sources are not perfect enough to be used or integrated. In connection with this problem, the problem of the automated formation of ontology based on the analysis of texts in natural language in construction is quite relevant. A common criterion for the quality of an ontology is based on an assessment of the convenience and efficiency of work and the connection of cognitive processes with semantic ones with the aim of constructing and updating it in the field of construction. Assessment of constructed ontology based on automated tools is another separate complex task. The main difference between the study, the results of which are presented in this work, is a cognitive-semantic analysis based on category theory, mathematical logic and universal algebra, algebra of sets, cortege algebra and relational algebra, namely, the construction of an ontological dictionary of construction, ontological constructions in open languages representing knowledge in construction. The work includes classification and definition of information units and their sources. The features of the information resource in the construction industry have been determined.

Key words: construction; ontology; concept; thesaurus; information object; isomorphism; analysis; semantic model.