

УДК 682.05

А.В. Шевчук, к.т.н., с.н.с.
Державне підприємство Поліграфічний комбінат "Україна"
з виготовлення цінних паперів

ВИКОРИСТАННЯ ТЕОРІЇ ФОРМАЛЬНИХ ГРАМАТИК ДЛЯ РОЗВ'ЯЗУВАННЯ ЗАДАЧ ПОБУДОВИ ГРАФІЧНИХ ЗАСОБІВ ЗАХИСТУ ПОЛІГРАФІЧНОЇ ПРОДУКЦІЇ

Показана доцільність використання теорії формальних граматики та визначені фактори, що є важливими при побудові графічних захисних елементів цінних паперів та документів суворого обліку.

Серед засобів захисту документів суворого обліку та цінних паперів досить часто використовуються окремі графічні елементи дизайну [1–3]. У більшості випадків такі засоби захисту або образи не мають власного сюжету, а є тим чи іншим чином побудованими узорами. Це означає, що для формування таких образів доцільно використовувати засоби теорії формальних граматики, які допускають досить зручну графічну інтерпретацію, яка у більшості випадків призводить до формування образів типу узорів різного виду. При побудові таких графічних елементів використовуються конкретно задані граматики та графічні базиси, які визначають той чи інший тип інтерпретації окремих елементів граматичних конструкцій. Коротко проаналізуємо основні можливості формальних граматики та розглянемо доцільність їх використання при побудові графічних засобів захисту.

У загальному випадку, формальна граматика записується у вигляді:

$$G = (T, V, S, P),$$

де T – скінченний алфавіт термінальних символів;

V – скінченний алфавіт нетермінальних символів;

S – початковий нетермінальний символ або аксіома;

P – скінченна система підстановок, ліві і праві частини яких є символами в скінченному алфавіті нетермінальних символів V .

При використанні засобів захисту у вигляді графічних образів типу узору чи орнаменту важливими є наступні фактори:

– такі зображення повинні бути ідентичними для кожного документа, з точністю до кожної компоненти орнаменту, і ця ідентичність не повинна забезпечуватись чисто технологічними засобами;

– алгоритм побудови орнаменту повинен бути детермінованим і забезпечувати єдиність розв'язку задачі побудови при заданих умовах;

– зміни вхідних даних функціонування алгоритму повинні приводити до змін конкретних фрагментів;

– кількість змін в орнаменті повинна визначатися однозначно змінами у вхідних даних алгоритму;

– обернена задача повинна належати до класу задач NP , чи P складності.

Розглянемо коротко на якісному рівні наведені фактори і покажемо, як вони забезпечуються у випадку використання формалізму теорії граматики. Слід відмітити, що формальні граматики розглядаються і тісно пов'язуються з теорією абстрактних автоматів, оскільки останні є, по суті, інтерпретаторами алгоритмів, що в рамках формальних граматики здійснюють ті чи інші перетворення, які, як правило, стосуються певних мов. У зв'язку з цим розглянемо уявлення про абстрактний автомат. Абстрактним автоматом згідно з [4], [5] називається система $U = (A, Q, B, \varphi, \psi)$,

де A, Q, B – скінченні алфавіти, які називаються вхідним алфавітом, множиною стану і вихідним алфавітом, відповідно;

φ – функція переходів;

ψ – функція виходів.

Зв'язок між формальними мовами та теорією формальних автоматів полягає у наступному. Нехай існує мова $L(G)$, яка задана на граматиці G . Це означає, що в граматиці G , у

відповідності до правил F^* , виділена підмножина слів $L(G)$. У формальних граматиках правила побудови слів підпорядковуються певним способам підстановок. Прототипом термінальної множини в теорії формальних граматик є множина станів автомата та множина вхідних та вихідних термів, з яких складаються відповідно вхідні та вихідні слова. Правила переходу автомата з одного стану в інший під дією вхідних слів представляють собою процедури виводу нового слова, яке з'являється на виході. Оскільки автомат є алгоритмічною структурою, то використовувати уявлення про автомат для моделювання процедур виводу слів досить зручно. Покажемо, як, задавши автомат певного типу, можна отримати мову, яка, у нашому випадку, представляє собою множину слів, і в цій мові всі слова мають певне обмеження. Така ситуація представляє інтерес при побудові графічних образів типу узорів чи орнаментів, які мають певні, наперед задані параметри або особливості. Нехай маємо недетермінований скінченний автомат (НСА). В подальшому будемо розглядати, у більшості випадків, скінченні автомати, оскільки вони працюють із скінченними словами. Нехай НСА допускає послідовність x , якщо в процесі читання цієї послідовності символів можна вибрати хоча б одну послідовність переходів у наступний стан так, щоб прийти з початкового стану в один з допустимих станів, то така послідовність буде словом певної мови $L(G)$.

Можна спроектувати автомат, а відповідно і алгоритм, який буде задовольняти тим чи іншим обмеженням або умовам, які накладаються на мову $L(x)$, що продукується автоматом. Щоб продемонструвати, що дає ця можливість в рамках досліджуваних задач, необхідно розглянути інтерпретацію відповідної мови $L(x)$ в предметній області графічних образів. Як відомо, графічні образи, що використовуються для захисту документа, можуть представляти собою двовимірні чи тривимірні образи, які можуть виготовлятися, у першому випадку, поліграфічним способом, наприклад, у вигляді тангінних сіток чи з використанням об'ємного ефекту, при побудові, наприклад, голографічних захисних елементів. Оскільки будь-яка мова може інтерпретуватись як окрема стандартна для всього образу компонента, то можна стверджувати, що деякий узор описується сукупністю слів, що гарантується автоматом. Кожна літера такого слова в графічній інтерпретації може представляти собою типову графічну компоненту, яка ідентифікує той чи інший символ в залежності від своєї орієнтації в деякій базовій системі всього документа і може представляти собою елементарний відрізок прямої лінії чи деякої кривої лінії, чи, взагалі, деякий графічний примітив. У цьому випадку кожне з правил граматики, при його графічній інтерпретації, визначає положення, в якому повинен знаходитися наступний елемент графічного примітива, чи відрізок, якщо останній прийнято в якості такого примітива. Якщо за графічний примітив використовується один тип графічної компоненти, то різні символи відповідної мови означають різні можливі положення цієї графічної компоненти в генерованому слові, яке відображає фрагмент зображення.

Зупинимось на випадку, коли узор чи графічний образ представляє собою з'єднані між собою елементарні відрізки, що відрізняються один від одного своєю орієнтацією на площині. Розглянемо, в чому полягають можливі методи захисту документів, що містять відповідний узор. Один з таких методів ґрунтується на визначеності повної кількості траєкторій, що існують в рамках зображення і задаються певними параметрами засобів побудови. До таких параметрів відносяться довжина траєкторії, що вимірюється кількістю елементарних відрізків, площа, яку займає певна траєкторія, та локалізація траєкторії.

Оскільки розглядається випадок узору, який складається з прямих відрізків, що з'єднуються між собою, то розглянемо деякі формальні аспекти побудови таких зображень, які ґрунтуються на уявленнях про формальні граматики. Метрика середовища або геометрична база середовища, в якому будується образ, в рамках формальних граматик інтерпретується, як набір певних продукцій побудови слів, а кожний термальний символ визначає певним чином орієнтований відрізок. Побудуємо граматику для випадку, коли кількість термінальних символів дорівнює восьми, у відповідності до геометричної бази середовища, в якому передбачається будувати зображення. Тому для формальної граматики $G = (V, T, P, S)$ будемо мати $T = \{a_1, a_2, \dots, a_8\}$. Геометрична база середовища, в якому розглядається образ, може інтерпретуватись як просторова сітка, яка розширює для нашого випадку уявлення про формальні граматики. Така сітка визначає дискретизацію просторового розміщення елементів зображення. Для дискретизації простору формування зображення можуть використовуватись прямокутні сітки, косокутні сітки, сітки, елементами яких є різні правильні геометричні

фігури, та інші. Така сітка задається певними правилами продукцій граматики і є множиною термінальних символів G . Тому основне призначення такої сітки полягає у впорядкуванні правил розміщення окремих елементів зображення. Від вибраної базової елементарної фігури для побудови сітки залежать можливості побудови більш чи менш складного зображення. Наприклад, коли вибрано в якості такого примітива квадрат чи ромб, то можна говорити про використання восьми термінальних символів, якщо вибрано в якості геометричного примітива правильний шестикутник, то кількість термінальних символів може бути рівною дванадцяти і т.д.

При виводі тих чи інших слів з граматики G , при створенні мови $L(G)$, в загальному правила продукції вибираються довільно або у відповідності до встановлених принципів, наприклад, вибирається те правило, яке призводить до повного вичерпання можливих підстановок. При використанні уявлень про абстрактні автомати, обмеження у використанні тих чи інших продукцій задаються у вигляді допустимих переходів автомата з одного стану в інший. У даному випадку, використання уявлень про абстрактний автомат може бути не достатньо ефективним, оскільки йдеться про генерацію слів, що відповідають фрагментам зображень, виходячи з початкового, стартового елемента. Тому інтерпретувати посилку редукції, як вхідне слово автомата, а вибір необхідного правила підстановки здійснювати на основі правил переходів з одного стану автомата в інший не обов'язково. У даному випадку введемо уявлення про правила генерації мови $L(G)$ з граматики G на основі використання системи правил підстановок P_i , яка входить у визначення граматики G , та правил вибору чергової продукції або правил управління генерацією компонент мови $L(G)$.

Розглянемо основні принципи формування таких правил, які, на відміну від правил продукцій P_i , будемо позначати символом R_i . Як і в кожному випадку, введення таких правил $r_i \in R$ повинно опиратись на деяку інтерпретацію, що може розглядатися, як модель, що має місце у відповідному випадку. Такою інтерпретацією є уявлення про геометричний базис середовища, в якому передбачається будувати графічний образ, або у випадку площини, на якій будемо будувати відповідний образ, таким базисом буде сітка, складена з геометрично правильних фігур, які використовуються як геометричні примітиви. Очевидно, що це не визначає необхідності нанесення такої сітки на відповідну поверхню або простір, в якому передбачається будувати зображення. Нехай за графічний примітив використовується прямокутник. Оскільки розмір прямокутника може вибиратися різним, від чого залежить технологічна густина зображення, що наноситься на поверхню, то у даному випадку в рамках формального підходу будемо приймати його довільним за розмірами і цей розмір приймати до уваги не будемо. Під параметром зображення, що уже згадувався вище, як густина образу будемо розуміти густину алгоритмічну. Під алгоритмічною густиною будемо розуміти кількість елементів зображення на одиницю умовної площі, що запишемо у вигляді $\gamma = m / S_U$. Умовну площу S_U будемо вимірювати кількістю графічних примітивів, які займають певну площу документа.

Таким чином, логічна густина образу не залежить від технологічної густини, яка визначається, наряду з логічною густиною образу, ще й абсолютними розмірами геометричних примітивів, в рамках яких будується відповідний геометричний образ. Слід підкреслити, що геометричним елементом образу не мусить бути певним чином орієнтований відрізок. Ним може бути довільний криволінійний елемент з певною, також заданою орієнтацією і, очевидно, зв'язаним з ним розміщенням в межах окремого фрагменту документа, який може задаватись аналогічною сіткою, що визначає базову метрику на площині документа, що підлягає захисту.

Розглянемо, виходячи з параметрів, що характеризують графічний образ, як можна формувати систему правил R_i , для вибору послідовних продукцій. До таких параметрів відносяться наступні характеристики зображення:

- повна кількість траєкторій, що існує в рамках графічного образу або повинна в ньому існувати (m);
- довжина траєкторії (d_i);
- площа, яку займає певна траєкторія (S_i);
- локалізація траєкторії або густина зображення фрагмента образу, що формується даною траєкторією (γ_i);

– подібність або повторюваність структури окремих траєкторій, як мінімум двох, що порівнюються (ψ);

– зв'язки між траєкторіями, які передбачають спільні точки початку траєкторій, точки перетину, що обумовлені функціонально, або інші зв'язки, що описуються тим чи іншим способом.

Повна кількість траєкторій, у нашому випадку, буде визначатися кількістю стартових символів або слів, які задаються в граматику G . Довжина траєкторії задається безпосередньо, оскільки формування слова, що відповідає одній траєкторії i , згідно з правилами R_i , які в подальшому будемо називати метаправилами, потребує інформації про довжину відповідного слова v_i .

Площа, яку займає певна траєкторія, вимірюється кількістю графічних примітивів, які вміщують елементарні графічні компоненти. Очевидно, що ця площа не відповідає величині площі реального документа. Площа, яку займає окрема траєкторія t_i , породжена стартовим символом $S_i(n_i)$, може співпадати з довжиною відповідної траєкторії, якщо кожний елемент траєкторії займає один графічний примітив. Очевидно, що у цьому випадку густина однієї траєкторії є найменша і дорівнює одиниці, або $\gamma = 1$. При певному виборі P_i з P можна досягнути заданої густини. Для кожного графічного примітива існує максимально можлива густина, яку в подальшому будемо називати густиною насичення (γ_n). Наприклад, якщо в якості графічного примітива вибрано прямокутник чи ромб, то густина насичення визначається максимальною кількістю елементарних відрізків, які можуть розміститися на даному примітиві. На прямокутнику таких різних за орієнтацією відрізків, при умові, що відрізки з'єднуються між собою і орієнтуються тільки по осях, що проходять через вершини, дорівнює шести. Таким чином, густина насичення в такій матриці простору буде дорівнювати $\gamma_n = 6$. Слід відмітити, що такий параметр, як густина насичення, є універсальним з точки зору вибору абсолютних величин графічних примітивів (їх розмірів). Оскільки кожний графічний примітив вміщає максимально допустиму кількість графічних елементів, то незалежно від вибраного способу розміщення таких елементів їх повинно бути не більше заданої максимальної кількості, яка визначається густиною насичення відповідного графічного примітива. Це зауваження стосується випадку, коли за графічну компоненту образу вибирається довільний криволінійний відрізок. У цьому випадку кількість способів його розміщення в графічному примітиві повинно обмежуватись величиною, що призводить до густини насичення γ_n . Якщо таких елементів розміщується більше в межах одного примітива, то абсолютні розміри графічного примітива необхідно зменшити. Випадки, коли окремі графічні компоненти не з'єднуються між собою при побудові траєкторій графічного образу, в даній роботі не розглядалися.

Подібність між двома фрагментами траєкторії χ_i є важливою характеристикою, оскільки подібність забезпечує певну симетрію графічному образу і зменшує його несиметричність. При дослідженні стійкості графічних образів до атак, або, при оцінці рівня захисту, який забезпечує той чи інший графічний образ, буде показано, що наявність симетрії χ_i в образах понижує рівень забезпечуваної ними міри захисту. Інтуїтивно це зрозуміло, якщо графічний образ має багато симетричних фрагментів, то така поширена атака, як фальсифікація документа здійснюється легше, ніж у випадку, коли графічний образ не має симетричних фрагментів. Перш за все, треба відмітити, що симетричні фрагменти можуть мати одну траєкторію, або одне слово, що породжується граматиною G , або дві окремі траєкторії t_i і t_j , що породжені однією граматиною G . Очевидно, що мати можливість управляти мірою подібності фрагментів ϕ_i і ϕ_k необхідно. Один зі способів такого управління може полягати у наступному. Якщо ми мали б можливість вводити певні умови або обмеження в граматику G , виходячи з потрібних значень величини χ_i , то таке управління мірою подібності або ступенем несиметричності було б можливим.

Розглянемо на якісному рівні можливі підходи до розв'язання цієї задачі. Розглянемо ознаки подібності окремих фрагментів різних траєкторій t_i і t_j . Геометрично, при аналізі самих образів, подібність χ_{ij} і χ_{kq} визначається як співпадання орієнтацій поточних компонент графічного образу у двох фрагментах, що порівнюються. З точки зору граматики G , це означає, що на деякому етапі формування слів почали використовуватись для двох окремих слів l_i і l_k однакові послідовності продукцій $\{P_i, \dots, P_{i+j}\}$. У цьому випадку управління цим процесом і, відповідно, забезпечення певної міри несиметричності повинно здійснюватись

метаправилами $r_i \in R$, які при виборі чергових $P_i \in P$ аналізують повторюваність послідовностей P_i, \dots, P_{i+j} .

Таким чином, проведені дослідження дозволили сформулювати основні параметри, що характеризують графічний образ. Підтверджена доцільність використання теорії формальних граматики для розв'язання задач побудови графічних елементів захисту друкованої продукції.

ЛІТЕРАТУРА:

1. Шевчук А.В. Методи формування параметрів графічних захищаючих зображень // Збірник наукових праць Інституту проблем моделювання в енергетиці НАН України. – Вип.17. – К., 2002. – С. 174–180.
2. Музика В.П., Шевчук А.В. Система захисту цінних паперів та документів суворого обліку – наукоємна проблема державного масштабу // Друкарство. – 2002. – № 4 (45). – С. 72–74.
3. Шевчук А.В. Технологічні проблеми виготовлення вітчизняних паспортів // Технологія і техніка друкарства: Збірник наукових праць. – Національний технічний університет (КПІ). – Київ, 2003. – Вип. 1. – С. 46–50.
4. Кудрявцев В.Б., Подколзин А.С., Ушчумлих Ш. Введение в теорию абстрактных автоматов. – М: МГУ, 1985. – 173 с.
5. Гладкий А.В. Формальные грамматики и языки. – М.: Наука, 1973.

ШЕВЧУК Анатолій Васильович – кандидат технічних наук, старший науковий співробітник, директор Державного підприємства Поліграфічний комбінат “Україна” з виготовлення цінних паперів, (Київ).

Наукові інтереси:

- інформаційні системи в поліграфії;
- розробка систем захисту цінних паперів та документів суворого обліку.

Тел.: (044) 213-98-90.

Подано 11.03.2004