

PODSTAWOWE ZASADY MODELOWANIA PROCESU PRODUKCJI ROLNICZEJ

Piotr Maksym

Katedra Podstaw Techniki, Uniwersytet Przyrodniczy w Lublinie

Streszczenie. W artykule przedstawione zostały podstawowe zasady dotyczące modelowania procesu produkcji rolniczej. Opisano podstawową fazę dotyczącą konceptualizacji dziedziny przedmiotowej oraz podstawowe założenia dotyczące jej formalizacji i wykonywalności w technologii sieci bayesowskich.

Słowa kluczowe: modelowanie procesu produkcji, konceptualizacja, sieci bayesowskie, inżynieria wiedzy

Wprowadzenie

Reprezentacja wiedzy w postaci modeli komputerowych, które są wbudowane w informatyczną infrastrukturę zarządzania procesami jest jednym z paradygmatów systemów produkcji opartych na wiedzy i informacji. Zdefiniowanie rolniczego procesu produkcyjnego, w postaci implementowanego modelu pojęciowego, musi uwzględniać wymagania jakościowe odnośnie produktu, jego bezpieczeństwo, jak i zagrożenia ekologiczne.

Zbudowanie modelu pojęciowego wymagało zdefiniowania pojęć i relacji międzypojęciowych dla standardowego procesu. Standard ten jest odniesiony do aktualnego stanu wiedzy o procesie i z założeniem, że działalność produkcyjna jest ulokowana na obszarach, które spełniają wymagania przyrodnicze i klimatyczne.

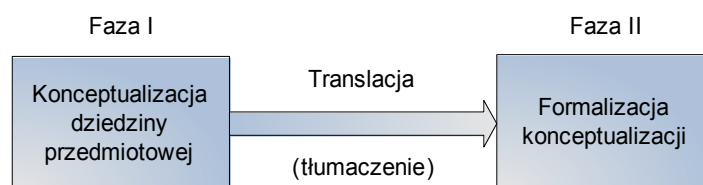
Metodyka modelowania komputerowych procesów produkcyjnych oparta na inżynierii wiedzy zawiera pojęcie „racjonalnego podmiotu działania” – jednego lub wielu. Zasada racjonalności oznacza podejmowanie decyzji wyłącznie w oparciu o informację o aktualnej sytuacji problemowej, możliwych do wykonania działaniach oraz ich oczekiwanej użyteczności. Informacje te są wynikiem ukierunkowanej percepcji (pomiarów) oraz racjonalnego rozumowania, w tym diagnostycznego i predykcyjnego [Marciniak 2005; Maksym i in. 2006]. Takim podmiotem działania jest w istocie każdy aktywny uczestnik procesu produkcyjnego. W literaturze z zakresu komputerowego modelowania procesów takich aktywnych uczestników procesu określa się mianem agentów. W przypadku typowych rolniczych procesów produkcyjnych podmiotem działania jest człowiek, ale też nie należy wykluczać sensowności takiego traktowania uprawianej rośliny, a także każdego biologicznego uczestnika procesu. Jest to przypadek systemu wieloagentowego, przy czym „agenci” ci nie koniecznie ze sobą kooperują, ale również rywalizują. Ponadto, mamy tu do czynienia z bardzo złożonym środowiskiem problemowym – częściowo obserwowalnym i stochastycznym, a więc niepewnym [Halpern 2003].

Cel i założenia wstępne pracy

Modele traktowane są jako symboliczne, formalne i wykonywalne reprezentacje wiedzy o modelowanej dziedzinie przedmiotowej. Niepewność jest związana z tą właśnie wiedzą, wynika z jej natury. W skonstruowanych modelach (sieciach bayesowskich) jako miarę niepewności zastosowano prawdopodobieństwo. W systemach reprezentacji wiedzy opartych na logice stosowane są również inne niż probabilistyczne miary niepewności. Szczególnie interesujące dla zastosowań inżynierskich są miary określane jako „plausibility measures” [Halpern 2003].

Celem pracy jest opracowanie metody modelowania rolniczego procesu produkcyjnego, uwzględniającej zasady metodyczne inżynierii wiedzy.

Proces modelowania składał się z dwóch faz (rys. 1.). Pierwsza polegała na nieformalnej konceptualizacji dziedziny przedmiotowej, którą jest proces uprawy pszenicy ozimej, w języku naturalnym, czyli utworzenia modelu pojęciowego rolniczego procesu produkcyjnego. Model pojęciowy rozumiany jest jako zbiór pojęć jakimi posługujemy się przy opisie procesu oraz relacji międzypojęciowych.



Źródło: opracowanie własne

Rys. 1. Ogólny schemat przedstawiający fazy procesu modelowania

Fig. 1. Main diagram showing modeling process phases

Druga faza to formalizacja konceptualizacji poprzez zdefiniowanie uniwersalnych sformalizowanych schematów pojęciowych. Wyróżniono tu trzy podstawowe schematy: kontrolowalny proces produkcji jako cykl kognitywny, problem rozpoznawania sytuacji/zagrożeń oraz problem predykcji. Użyte schematy pojęciowe ustrukturalizowały zgromadzoną wiedzę, co pozwoliło na zredukowanie złożoności modeli.

Podstawowy problem metodyczny dotyczył algorytmizacji konstruowania sieci bayesowskich. Zadanie to sprowadzono do hierarchicznej modularyzacji sieci, tj. tworzenia kilku warstw modułów zaczynając od najbardziej ogólnych i stopniowo schodząc do coraz to bardziej szczegółowych. Wyodrębniono dwie kategorie modułów, odpowiadające dwóm typowym mechanizmom rozumowania – rozpoznawania i predykcji. Warstwa modułów szczegółowych odpowiada różnym typom procesów uprawowych i specyficznym stadiom rozwojowym roślin.

Najważniejszymi z opracowanych modułów funkcjonalnych są: moduł rozpoznawania fazy rozwojowej, moduł rozpoznawania zagrożeń, moduł predykcji plonu oraz moduł fuzji heterogenicznych informacji. Wybór odpowiedniej technologii modelowania wynikał z potrzeby uwzględnienia w modelach czynnika niepewności, co jest niezbędne w mode-

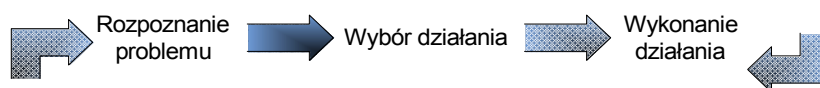
lach wspomagających i racjonalizujących zarządzanie ryzykiem w produkcji rolniczej. Technologia sieci bayesowskich umożliwia budowanie modeli poprzez zastosowanie algorytmów uczenia maszynowego, a następnie ich wykorzystywanie jako zasobu produkcyjnego trwale wbudowanego w proces produkcyjny i podlegającego adaptacji do modelowanego środowiska, w oparciu o efektywne algorytmy inferencyjne [Conrady, Jouffe 2011].

Do zbudowania i walidacji modeli użyto profesjonalnego systemu modelowania BayesiaLab. Użycie adekwatnych i efektywnych narzędzi modelowania było obok dobrej, głębokiej i całościowej znajomości przedmiotu modelowania pierwszym krokiem do otrzymania adekwatnych, wystarczająco dokładnych modeli. System BayesiaLab posiada takie możliwości, między innymi dlatego, że wymusza i wspiera tworzenie uporządkowanej i spójnej konceptualizacji przedmiotu modelowania oraz częściowo standaryzuje proces modelowania tak, że jego produkty (modele) są w znacznym stopniu powtarzalne, przewidywalne i porównywalne. Ekspresywność języka sieci bayesowskich jest na tyle bogata, że modele mogą być wystarczająco dobre [Kusz, Marciniak 2006].

System powinien być adaptacyjny i funkcjonować jak baza wiedzy. Działanie bazy wiedzy polega na zadawaniu pytań i generowaniu odpowiedzi w zależności od uzyskanych od użytkownika informacji. Formułowanie pytań polega na specyfikowaniu określonej wartości („twardy fakt”) lub rozkładu prawdopodobieństwa („miękki fakt”) nad określonymi zmiennymi. Pytania mogą mieć charakter predykcyjny lub wyjaśniający. Produkowanie odpowiedzi na zadawane pytania odbywa się poprzez obliczenie nowych rozkładów prawdopodobieństwa nad innymi zmiennymi. W odróżnieniu od bazy danych, w której wynikiem jest rezultat przeszukiwania według określonych kryteriów, baza wiedzy oblicza wynik, tworząc nowe rozkłady prawdopodobieństwa nad zmiennymi.

Metodyka

U podstaw założeń metodycznych leżało rozumienie modelowania dziedziny przedmiotowej jako procesu translacji wiedzy wyrażonej w języku naturalnym do postaci kodu, który jest formalny i wykonywalny. Samo przekształcanie wiedzy z reprezentacji nieformalnej w formalną odbywało się za pomocą ekspresywnych schematów pojęciowych. Schematy te były podstawą do modularyzacji systemu. Jako przykładowy schemat przedstawiono cykl kognitywny procesu decyzyjnego (rys. 2.), tzn. działania inicjujące oraz wszystkie celowe ingerencje człowieka (podmiotu zmieniającego stan rzeczywistości fizycznej) w proces produkcyjny [Maksym i in. 2006].



Źródło: opracowanie własne

Rys. 2. Schemat cyklu kognitywnego procesu decyzyjnego
Fig. 2. Diagram of cognitive cycle of the decision process

Należy posiadać dostęp do źródeł informacji (np.: obserwacje, sensory), następnie informacje te trzeba przekształcić, aby dowiedzieć się jakie są nieprawidłowości i jak te nieprawidłowości można usunąć czy też zniwelować. Procesy, które przebiegają według przedstawionego cyklu kognitywnego są procesami opartymi na wiedzy i informacji.

Cykl kognitywny opiera się na dwóch subschematach:

- schemacie rozumowania dedukcyjnego, który służy do predykcji, co oznacza, że wszystkie decyzje i wynikające z nich działania, muszą brać pod uwagę, zarówno informacje o aktualnym stanie procesu, jak i predykcję skutków tych działań, które podejmujemy;
- schemacie rozumowania hipotetyczno-dedukcyjnego, który służy do rozpoznawania (diagnozowania) aktualnego stanu procesu w oparciu o obserwowalne jego atrybuty (informacja najczęściej niekompletna).

Modelowanie procesu polega na podejmowaniu decyzji statystycznych oraz ich implementowaniu w języku sieci bayesowskich. Sieć bayesowska jest acyklicznym grafem skierowanym (ang. *DAG – DirectedAcyclicGraph*), który składa się z dwóch podstawowych elementów: węzłów, które reprezentują zmienne losowe oraz łuków łączących węzły, które reprezentują probabilistyczne zależności, w tym przyczynowo-skutkowe między tymi zmiennymi. Z węzłami związane są tabele warunkowych rozkładów prawdopodobieństwa (ang. *CPT – ConditionalProbabilityTable*) [Pearl 1988]. Należy zwrócić uwagę, że liczba pól do wypełnienia rośnie wykładniczo i zależy od liczby węzłów poprzedzających-wstępujących oraz od liczby stanów w poszczególnych węzłach. Każda zmienna ma ściśle określony typ i dziedzinę wartości. Źródłem danych do wyznaczenia warunkowego rozkładu prawdopodobieństwa mogą być dane empiryczne lub wiedza, którą posiada ekspert - praktyk.

Podsumowanie

Budowany komputerowy model procesu produkcji pszenicy ozimej będzie służył do wspomagania decyzji w zarządzaniu tym procesem.

Metodyka budowy wykonywalnych i formalnych modeli oparta jest na aktualnych osiągnięciach inżynierii wiedzy, które wykorzystują technologię sieci bayesowskich jako formalny i wykonywalny system reprezentacji wiedzy. Język ten tworzy uniwersalny system dla specyfikacji, konceptualizacji i analizy sytuacji problemowych w warunkach nie-deterministycznych. Zaprezentowane podejście do modelowania procesu rolniczego z wykorzystaniem cyklu kognitywnego bardzo dobrze wpisuje się w schemat integrowanej produkcji.

Tworzony model komputerowy zakłada implementację za pomocą sieci bayesowskich i funkcjonowanie jako bazy wiedzy. Baza ta za pomocą zgromadzonego zbioru faktów oraz mechanizmów wnioskowania odpowiadać będzie na zadawane pytania. Mechanizmy uczenia i wnioskowania wbudowane w sieci bayesowskie zapewniają adaptacyjność modelu, na przykład do oceny zagrożeń w różnych fazach rozwojowych rośliny. Użyte schematy pojęciowe umożliwiły modularyzację komputerowego modelu procesu produkcyjnego, co pozwala na zautomatyzowanie procesu uczenia zarówno na poziomie warunkowych rozkładów prawdopodobieństw, jak i na poziomie topologii sieci.

Bibliografia

- Conrady S., Jouffe L.** 2011. Breast Cancer Diagnostics with Bayesian Networks [online]. [dostęp 12.05.2011]. Dostępny w internecie http://www.conradyscience.com/white_papers/WBCD_V7.pdf
- Halpern J.Y.** 2003. Reasoning about uncertainty. MIT Press. ISBN 9780262083201. s. 11-68, 189-237.
- Kusz A., Marciniak A.** 2006. Dynamiczne sieci probabilistyczne jako system reprezentacji wiedzy. Inżynieria Rolnicza. Nr 12(87). s. 285-294.
- Maksym P., Marciniak A., Kostecki R.** 2006. Zastosowanie sieci bayesowskich do modelowania rolniczego procesu produkcyjnego. Inżynieria Rolnicza. Nr 12(87). s. 321-330.
- Marciniak A.** 2005. Projektowanie systemu reprezentacji wiedzy o rolniczym procesie produkcyjnym. Rozprawa naukowa. Z. 298, Wydawnictwo AR Lublin. ISSN 0860-4355.
- Pearl J.** 1988. Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers. ISBN 9781558604797. s. 1-27.

BASIC RULES OF AGRICULTURAL PRODUCTION PROCESS MODELLING

Abstract. The article presents basic rules of agricultural production process modeling. It describes the main phase of field conceptualization and the main guidelines for their formalization and executability in technology of Bayesian networks.

Key words: modeling of production process, conceptualization, Bayesian networks, knowledge engineering

Adres do korespondencji:

Piotr Maksym; e-mail: piotr.maksym@up.lublin.pl
Katedra Podstaw Techniki
Uniwersytet Przyrodniczy w Lublinie
ul. Doświadczalna 50a
20-280 Lublin