XV Krajowa Konferencja Automatyki

Tom I

Redaktorzy:
Zdzisław Bubnicki
Roman Kulikowski
Janusz Kacprzyk

XV Krajowa
Konferencja Automatyki
Tom I

KKA 2005

Redaktorzy:
Zdzisław BUBNICKI
Roman KULIKOWSKI
Janusz KACPRZYK

ORGANIZATOR
Komitet Automatyki i Robotyki Polskiej Akademii Nauk
Instytut Badań Systemowych Polskiej Akademii Nauk
WSPÓŁORGANIZATORZY
Politechnika Warszawska
Przemysłowy Instytut Automatyki i Pomiarów
Polskie Stowarzyszenie Pomiarów, Automatyki i Robotyki

ORGANIZATOR

Komitet Automatyki i Robotyki Polskiej Akademii Nauk
Instytut Badań Systemowych Polskiej Akademii Nauk

WSPÓŁORGANIZATORZY

Politechnika Warszawska
Przemysłowy Instytut Automatyki i Pomiarów
Polskie Stowarzyszenie Pomiarów, Automatyki i Robotyki

KOMITET PROGRAMOWY

Przewodniczący	Zdzisław BUBNICKI
Zastępca Przewodniczącego	Roman KULIKOWSKI

CZŁONKOWIE

Stanisław BAŃKA
Mikołaj BUSŁOWICZ
Ryszard GESSING
Jakub GUTENBAUM
Stanisław KACZANOWSKI
Janusz KACPRZYK
Józef KORBICZ
Krzysztof KOZŁOWSKI
Krzysztof KUŹMIŃSKI
Krzysztof MALINOWSKI
Antoni NIEDERLIŃSKI
Tadeusz PUCHAŁKA
Stanisław SKOCZOWSKI
Jerzy ŚWIĄTEK
Ryszard TADEUSIEWICZ
Krzysztof TCHOŃ
Jan WĘGLARZ
Michał BIAŁKO
Władysław FINDEISEN
Henryk GÓRECKI
Jerzy JÓZEFCZYK
Tadeusz KACZOREK
Jerzy KLAMKA
Zbigniew KOWALSKI
Juliusz L. KULIKOWSKI
Kazimierz MALANOWSKI
Wojciech MITKOWSKI
Władysław PEŁCZEWSKI
Leszek RUTKOWSKI
Roman SŁOWIŃSKI
Andrzej ŚWIERNIAK
Piotr TATJEWSKI
Leszek TRYBUS
Andrzej P. WIERZBICKI

KOMITET ORGANIZACYJNY

Przewodniczący	Roman KULIKOWSKI
Zastępcy Przewodniczącego	Janusz KACPRZYK Stanisław KACZANOWSKI Tadeusz KACZOREK Krzysztof MALINOWSKI
Członkowie	Roman OSTROWSKI Tadeusz PUCHAŁKA Dariusz WAGNER
Sekretarze naukowi	Jan STUDZIŃSKI Jan W. OWSIŃSKI

ISBN 83-89475-00-6

Druk: ARGRAF, Warszawa

REFERATY PLENARNE

WIELOETAPOWE STEROWANIE ROZMYTE ZE STOCHASTYCZNYM UKŁADEM STEROWANYM

Janusz KACPRZYK

Instytut Badań Systemowych Polskiej Akademii Nauk
ul. Newelska 6, 01-447 Warszawa, e-mail:kacprzyk@ibspan.waw.pl

Streszczenie: Rozpatruje się wieloetapowe sterowanie rozmyte łańcuchem Markowa z rozmytymi ograniczeniami na sterowania i rozmytymi celami rozmytymi nałożonymi na osiągane stany, ze skończonym i nieskończonym horyzontem planowania. Analizuje się możliwość użycia różnych operatorów agregacji w funkcji celu (decyzji rozmytej). Pokazuje się zastosowanie do planowania rozwoju regionalnego. Wskazuje się na możliwe rozszerzenia, m.in. poprzez użycie zmiennych niepewnych Bubnickiego.

Słowa kluczowe: wieloetapowe sterowanie rozmyte, rozmyte programowanie dynamiczne, operatory agregacji, planowanie regionalne, horyzont planowania

1. WPROWADZENIE

Rozpatrujemy zadanie wyznaczania optymalnych sterowań w zadaniu wieloetapowego sterowania rozmytego (por. Kacprzyk [23]), na skończonym i nieskończonym horyzoncie, gdy układ sterowany jest stochastyczny (łańcuch Markowa). Do rozwiązywania stosujemy głównie programowanie dynamiczne dla horyzontu skończonego, a metody typu iteracji polityk dla horyzontu nieskończonego. Pokazujemy przykład zastosowania do planowania tzw. zrównoważonego rozwoju regionalnego. Wskazujemy też na możliwość rozszerzenia modeli m.in. przez użycie aparatu zmiennych niepewnych Bubnickiego [5], [6].

Punktem wyjścia jest programowanie dynamiczne (por. Bellman [1], [3], [8], a przede wszystkim Sniedowich [31]). Możliwość zastosowania aparatu zbiorów rozmytych Zadeha[35] w programowaniu dynamicznym pokazali Bellman i Zadeh [2], a potem m.in. Kacprzyk [17], [18], [22], [23], co doprowadziło do powstania tzw. rozmytego programowania dynamicznego. Te modele znalazły wiele zastosowań (patrz np. książki Kacprzyka [22], [23]), ale tu wspomnimy głównie o zastosowaniu do planowania zrównoważonego rozwoju regionalnego, co zaproponowali Kacprzyk i Straszak [29] – por. Kacprzyk [24].

2. Zbiory rozmyte i rozmyte układy dynamiczne

Teoria zbiorów rozmytych Zadeha [35] jest prostym, ale silnym i efektywnym narzędziem do reprezentacji i przetwarzania informacji nieprecyzyjnej typu "*wysoki* budynek", "*duża liczba*" itp.

Zbiór rozmyty A w przestrzeni rozważań $X = \{x\}$ określa się jako zbiór par $A = \{(\mu_A(x), x)\}$, gdzie $\mu_A : X \longrightarrow [0,1]$ jest *funkcją przynależności* zbioru rozmytego A, a $\mu_A(x) \in [0,1]$ jest *stopniem przynależności* elementu $x \in X$ w zbiorze rozmytym A. W zbiorze rozmytym przejście od przynaleńości do nieprzynależności elementu do zbioru jest stopniowe (od 0 do 1), a nie skokowe (albo 0 albo 1) jak w zbiorze zwykłym. W dalszej cześci utożsamiać będziemy zbiory rozmyte z ich funkcjami przynależności, a także często z ich etykietami.

Jeżeli chodzi o podstawowe własności, to zbiór rozmyty A w X jest *pusty*, $A = \emptyset$, wtedy i tylko wtedy, gdy $\mu_A(x) = 0$, dla wszystkich $x \in X$. Dwa zbiory rozmyte A i B w X są *równe*, $A = B$, wtedy i tylko wtedy, gdy $\mu_A(x) = \mu_B(x)$, dla wszystkich $x \in X$. Zbiór rozmyty A w X jest *zawarty w* lub jest *podzbiorem* B w X, $A \subseteq B$, wtedy i tylko wtedy, gdy $\mu_A(x) \leq \mu_B(x)$, dla wszystkich $x \in X$.

Dopełnienie zbioru rozmytego A w X, $\neg A$, odpowiadające negacji "nie", określamy jako

$$\mu_{\neg A}(x) = 1 - \mu_A(x); \forall x \in X \tag{1}$$

Przecięcie dwu zbiorów rozmytych A i B w X, $A \cap B$, odpowiadające spójnikowi "i", określamy jako

$$\begin{aligned}\mu_{A\cap B}(x) &= \min[\mu_A(x), \mu_B(x)] = \\ &= \mu_A(x) \wedge \mu_B(x); \forall x \in X\end{aligned} \tag{2}$$

Sumę dwu zbiorów rozmytych A i B w X, $A + B$, odpowiadającą spójnikowi "lub", definiujemy jako

$$\begin{aligned}\mu_{A+B}(x) &= \max[\mu_A(x), \mu_B(x)] = \\ &= \mu_A(x) \vee \mu_B(x); \forall x \in X\end{aligned} \tag{3}$$

Te tradycyjne definicje przecięcia i sumy można uogólnić stosując, odpowiednio, tzw. t-normy i s-normy (t-konormy).

t-normą jest funkcja $t : [0,1] \times [0,1] \longrightarrow [0,1]$, taką że, dla każdego $a, b, c \in [0,1]$: (1) $t(a,1) = a$, (2) $a \leq b \Longrightarrow t(a,c) \leq t(b,c)$, (3) $t(a,b) = t(b,a)$, (4) $t[a, t(b,c)] = t[t(a,b), c]$. Oczywiście, t-norma jest monotonicznie niemalejąca względem obu argumentów, a ponadto $t(a,0) = 0$. Ważniejszymi t-normami są: (1) $t(a,b) = a \wedge b = \min(a,b)$, (2) $t(a,b) = a \cdot b$, (3) t-norma Łukasiewicza, czyli $t(a,b) = \max(0, a + b - 1)$.

s-norma (lub *t-konorma*) to funkcja $s : [0,1,] \times [0,1] \longrightarrow [0,1]$, taką że, dla każdego $a,b,c \in [0,1]$: (1) $s(a,0) = a$, (2) $a \leq b \Longrightarrow s(a,c) \leq s(b,c)$, (3) $s(a,b) = s(b,a)$, (4) $s[a,s(b,c)] = s[s(a,b),c]$. Ważniejszych s-normami są: (1) $s(a,b) = a \vee b = \max(a,b)$, (2) $s(a,b) = a + b - ab$, (3) s-norma Łukasiewicza, czyli $s(a,b) = \min(a+b,1)$.

Relację rozmytą R między dwoma (nierozmytymi) zbiorami $X = \{x\}$ i $Y = \{y\}$ definiuje się jako zbiór rozmyty w $X \times Y$, tzn. $R = \{(\mu_R(x,y),(x,y))\} = \{\mu_R(x,y)/(x,y)\}$, dla każdego $(x,y) \in X \times Y$, gdzie $\mu_R(x,y) : X \times Y \longrightarrow [0,1]$ jest funkcją przynależności relacji rozmytej R, a $\mu_R(x,y) \in [0,1]$ jest stopniem, w którym elementy $x \in X$ i $y \in Y$ są ze sobą w relacji R. Tę definicję można oczywiście rozszerzyć na k-narne relacje rozmyte. W tym kontekscie zbiór rozmyty jest unarną relacją rozmytą. Ponieważ relacja rozmyta jest zbiorem rozmytym, więc wszystkie definicje, właściwości, operacje itp. na zbiorach rozmytych podane powyżej przenoszą się na relacje rozmyte.

Pierwsze podejście do określenia prawdopodobieństwa zdarzeń rozmytych, np. gdy pytamy o to, jakie jest prawdopodobieństwo, że jutro będzie *dobra pogoda*, pochodzi od Zadeha [36]. Zdarzenie rozmyte jest zbiorem rozmytym A w $X = \{x\} = \{x_1,\ldots,x_n\}$ o mierzalnej w sensie Borela funkcji przynależności. Prawdopodobieństwa (nierozmytych) zdarzeń elementarnych $x_1,\ldots,x_n \in X$ są znane i równe, odpowiednio, $p(x_1),\ldots,p(x_n) \in [0,1]; p(x_1) + \cdots + p(x_n) = 1$.

Prawdopodobieństwo (nierozmyte) zdarzenia rozmytego A w $X = \{x_1,\ldots,x_n\}$, $p(A)$, określa się jako

$$p(A) = \sum_{i=1}^{n} \mu_A(x_i)p(x_i) \tag{4}$$

tzn. jako wartość oczekiwaną funkcji przynależności zbioru rozmytego A, $\mu_A(x)$. Zauważmy, że: (1) $p(\emptyset) = 0$, (2) $p(\neg A) = 1 - p(A)$, (3) $p(A+B) = p(A) + p(B)$ - $p(A \cap B)$, itd.

Ta klasyczna definicja (nierozmytego) prawdopodobieństwa zdarzenia rozmytego jest najczęściej stosowana i będzie tu także stosowana.

Kluczowym elementem w naszym modelu wieloetapowego sterowania rozmytego jest *dynamiczny układ sterowany*, rozumiany jako określenia przejść stanów. Zakłada się tu, że taka zależność jest znana dla wszystkich etapów sterowania. Może ona być deterministyczna, stochastyczna lub rozmyta.

Niech *przestrzenią stanów* będzie $X = \{s_1,\ldots,s_n\}$, a *przestrzenią sterowań* – $U = \{c_1,\ldots,c_m\}$; z założenia obie będą skończone. Przejścia stanów *deterministycznego układu sterowanego* określa funkcja $f : X \times U \longrightarrow X$, taka że

$$x_{t+1} = f(x_t,u_t); t = 0,1,\ldots \tag{5}$$

gdzie $x_t, x_{t+1} \in X$ są stanami na etapach, opowiednio, t and $t+1$, a $u_t \in U$ jest sterowaniem na t.

Przejścia stanów *stochastycznego układu sterowanego* (łańcucha Markowa) określa funkcja $p : X \times U \times X \longrightarrow [0,1]$, taka że

$$p(x_{t+1} \mid x_t,u_t) \in [0,1]; t = 0,1,\ldots \tag{6}$$

jest prawdopodobieństwem warunkowym osiągnięcia $x_{t+1} \in X$ z $x_t \in X$ i przy przyłożeniu $u_t \in U$.

Dla układu rozmytego, stan rozmyty na etapie t jest określony jako zbiór rozmyty X_t określony w X, którego funkcją przynależności jest $\mu_{X_t}(x_t)$. Sterowanie na etapie t może być albo nierozmyte, tzn. $u_t \in U$, albo rozmyte, tzn. scharakteryzowane przez zbiór rozmyty U_t określony w U, którego funkcją przynależności jest $\mu_{U_t}(u_t)$. Przejścia stanów *rozmytego układu sterowanego* określa funkcja

$$X_{t+1} = F(X_t,U_t); t = 0,1,\ldots \tag{7}$$

3. Wieloetapowe sterowanie rozmyte

Punktem wyjścia jest ogólne podejście do podejmowania decyzji i sterowania w warunkach rozmytości pochodzące od Bellmana i Zadeha [2], w którym wprowadzono tzw. *otoczenia rozmytego*: cele rozmyte, ograniczenia rozmyte i decyzję rozmytą określoną w zbiorze możliwych *opcji* (alternatyw, wariantów, decyzji, ...), $X = \{x\}$.

Cel rozmyty definiuje się jako zbiór rozmyty G w X, $\mu_G : X \longrightarrow [0,1]$, *ograniczenie rozmyte* definiuje się podobnie jako zbiór rozmyty C w X, $\mu_C : X \longrightarrow [0,1]$. Ta identyczność traktowania G i C sugeruje następujące ogólne postawienie zadania starowania

"Osiągnąć G i spełnić C" (8)

co można wyrazić jako przecięcie zbiorów rozmytych:

$$\mu_D(x) = \\ = \mu_G(x) \wedge \mu_C(x) = \min[\mu_G(x),\mu_C(x)]; \forall x \in X \tag{9}$$

Jeżeli chcemy zaimplementować takie rozwiązanie rozmyte, to musimy znaleźć rozwiązanie nierozmyte, tzw. *decyzję maksymalizującą*, $x^* \in X$, taką że:

$$\mu_D(x^*) = \max_{x \in X} \mu_D(x) \tag{10}$$

Ważna jest możliwość uwzględnienia ważności poszczególnych ograniczeń i celów rozmytych, czy też etapów sterowania (tzw. dyskonto). Możliwe są tu różne podejścia, np. jeżeli $w_G, w_C \in [0,1]$ są ważnościami G i C, to:

$$\begin{aligned}
\mu_D(x) &= [(1-w_C) \vee \mu_C(x)] \wedge && (11)\\
&\wedge[(1-w_G) \vee \mu_G(x)] && (12)\\
\mu_D(x) &= [(1-w_C) + w_c\mu_C(x)] \wedge && (13)\\
&\wedge[(1-w_G) + w_g\mu_G(x)] && (14)\\
\mu_D(x) &= [\mu_C(x)]^{w_C} \wedge [\mu_G(x)]^{w_G} && (15)
\end{aligned}$$

W praktyce zwykle występuje wiele celów i ograniczeń rozmytych, $G_1,\ldots,G_n, C_1,\ldots,C_m$, i wtedy:

$$\mu_D(x) = \mu_{G_1}(x) \wedge \ldots \wedge \mu_{G_n}(x) \wedge \\ \wedge \mu_{C_1}(x) \wedge \ldots \wedge \mu_{C_m}(x); \forall x \in X \tag{16}$$

Zwykle C jest określone jako zbiór rozmyty w $X = \{x\}$, G jest określony jako zbiór rozmyty w $Y = \{y\}$ oraz dana jest funkcja $f : X \longrightarrow Y, y = f(x)$, a wtedy:

$$\mu_D(x) = \mu_G[f(x)] \wedge \mu_C(x); \forall x \in X \tag{17}$$

oraz dla $G_1, \ldots, G_n$ określonych w Y i $C_1, \ldots, C_m$ określonych w X, i $f: X \longrightarrow Y, y = f(x)$:

$$\mu_D(x) = \mu_{G_1}[f(x)] \wedge \cdots \wedge \mu_{G_n}[f(x)] \wedge \wedge \mu_{C_1}(x) \wedge \cdots \wedge \mu_{C_n}(x); \forall x \in X \quad (18)$$

a także możemy wprowadzić dyskontowanie.

We wszystkich przypadkach poszukujemy *decyzji maksymalizującej* $\mu_D(x^*) = \max_{x \in X} \mu_D(x)$.

3.1. Wieloetapowe sterowanie rozmyte

Przedstawimy najpierw prosty model wyjściowy, a potem pokażemy możliwe jego rozszerzenia. Istotę wieloetapowego sterowania w rozmytym otoczeniu (w warunkach rozmytości) pokazano na Rysunku 1.

Załóżmy, że przestrzenią sterowań jest $U = \{u\} = \{c_1, \ldots, c_m\}$, a przestrzenią stanów jest $X = \{x\} = \{s_1, \ldots, s_n\}$. Sterowania utożsamiać będziemy z wejściem, a stan z wyjściem. Załóżmy najpierw, że układ sterowany jest deterministyczny o przejściach stanów danych jako

$$x_{t+1} = f(x_t, u_t); t = 0, 1, \ldots \quad (19)$$

przy czym (19) reprezentuje stacjonarny deterministyczny układ sterowany, a w przypadku układu niestacjonarnego, mamy $f_t : X \times U \longrightarrow X$, takie że $x_{t+1} = f_t(x_t, u_t)$, $t = 0, 1, \ldots$.

Na każdym t, na $u_t \in U$ jest nałożone *ograniczenie rozmyte* $\mu_{C^t}(u_t)$, a na $x_{t+1} \in X$ jest nałożony *cel rozmyty* $\mu_{G^{t+1}}(x_{t+1})$; $t = 0, 1, \ldots$. *Czas zakończenia* (horyzont), tzn. maksymalną liczbę etapów, oznaczamy jako $N \in \{1, 2, \ldots\}$; może on być skończony albo nie.

Wskaźnikiem jakości jest decyzja rozmyta $D(x_0)$:

$$\mu_D(u_0, \ldots, u_{N-1} \mid x_0) = \mu_{C^0}(u_0) \wedge \mu_{G^1}(x_1) \wedge \wedge \ldots \wedge \mu_{C^{N-1}}(u_{N-1}) \wedge \mu_{G^N}(x_N) \quad (20)$$

gdzie kolejne $x_0, u_0, x_1, u_1, x_2, \ldots, x_{t-1}, u_{t-1}$ wyznaczone są przez równianie przejśc stanów (19).

Częściej stosuje się nieco uproszczoną postać decyzji rozmytej, zakładając, że na $u_0, u_1, \ldots, u_{N-1}$ nałożone są ograniczenia rozmyte, $\mu_{C^0}(u_0)$, $\mu_{C^1}(u_1)$, ..., $\mu_{C^{N-1}}(u_{N-1})$, natomiast cel rozmyty jest nałożony jedynie na stan końcowy x_N, $\mu_{G^N}(x_N)$. W takim przypadku:

$$\mu_D(u_0, \ldots, u_{N-1} \mid x_0) = \mu_{C^0}(u_0) \wedge \ldots \ldots \wedge \mu_{C^{N-1}}(u_{N-1}) \wedge \mu_{G^N}(x_N) \quad (21)$$

Ważna jest możliwość przypisania wag do poszczególnych t, $w_0, w_1, \ldots, w_{N-1} \in [0, 1]$, głównie wprowadzenie tzw. *dyskonta*: ważność tego, co ma miejsce wcześniej, jest większa niż tego, co później. Na przykład ($b > 1$):

$$\begin{aligned} \mu_D(u_0, \ldots, u_{N-1} \mid x_0) &= \\ &= b^0[\mu_{C^0}(u_0) \wedge \mu_{G^1}(x_1)] \wedge \ldots \\ &\ldots \wedge b^{N-1}[\mu_{C^{N-1}}(u_{N-1}) \wedge \mu_{G^N}(x_N)] = \\ &= \bigwedge_{t=0}^{N-1} b^t[\mu_{C^t}(u_t) \wedge \mu_{G^{t+1}}(x_{t+1})] \end{aligned} \quad (22)$$

Zadanie polega na znalezieniu optymalnego ciągu sterowań $u_0^*, \ldots, u_{N-1}^*$, $u_t^* \in U$, $t = 0, 1, \ldots, N-1$, takiego że:

$$\mu_D(u_0^*, \ldots, u_{N-1}^* \mid x_0) = = \max_{u_0, \ldots, u_{N-1} \in U} \mu_D(u_0, \ldots, u_{N-1} \mid x_0) \quad (23)$$

Dogodnie jest wyrazić sterowania w postaci *polityki sterowania* (polityki) dla etapu t: $a_t : X \longrightarrow U$, $t = 0, 1, \ldots$, takiej że $u_t = a_t(x_t)$, $t = 0, 1, \ldots$. Nie zawsze można jednak przyjąć taką prostą definicję polityki, a wtedy określić trzeba ją jako: $a_t : X \times U \times X \times \cdots \times U \times X \longrightarrow U$, $t = 0, 1, \ldots$, taką że $u_t = a_t(x_0, u_0, x_1, \ldots, u_{t-1}, x_t)$, $t = 0, 1, \ldots$. tzn. u_t zależy nie tylko od x_t, ale także od całej przeszłej trajektorii. W praktyce przyjmuje to zwykle prostszą postać $a_t : [0, 1] \times X \longrightarrow U$, $t = 0, 1, \ldots$, taką że $u_t = a_t[w_t(x_0, u_0, x_1, \ldots, u_{t-1}), x_t]$, $t = 0, 1, \ldots$. gdzie $w_t : X \times U \times X \times \cdots \times X \times U \longrightarrow [0, 1]$, $t = 0, 1, \ldots$, jest funkcją "podsumowującą" przeszłą trajektorię; tak więc, u_t zależy od x_t i od "podsumowania" przeszłości (trajektorii do $t-1$). Więcej informacji można znaleźć u Iwamoto [13], Iwamoto i Sniedovicha[14] lub Iwamoto, Tsurusakiego i Fujity [15], albo w książkach Kacprzyka [22],[23].

Ważne jest pojęcie *polityki stacjonarnej*, $a : X \longrightarrow U$, takiej że $u_t = a(x_t)$, $t = 0, 1 \ldots$. Czasami potrzebna jest polityka stacjonarna typu $a : X \times U \times X \times \cdots \times U \times X \longrightarrow U$, taka że $u_t = a(x_0, u_0, x_1, \ldots, u_{t-1}, x_t)$, $t = 0, 1, \ldots$, a często można także przyjąć uproszczoną postać $a : [0, 1] \times X \longrightarrow U$, taką że $u_t = a[w(x_0, u_0, x_1, \ldots, u_{t-1}), x_t]$, $t = 0, 1, \ldots$, gdzie $w : X \times U \times X \times \cdots \times X \times U \longrightarrow [0, 1]$, $t = 0, 1, \ldots$, jest funkcją "podsumowującą" przeszłą trajektorię.

Strategię sterowania (strategię) określa się jako ciąg polityk (sterowania) $A = (a_0, a_1, \ldots, a_{N-1})$, a *strategię stacjonarną* jako $a_N = a = \underbrace{(a, a, \ldots, a)}_{N}$. Oczywiście, dla nieskończonego czasu zakończenia strategię (właściwie strategię stacjonarną), oznacza się jako $a_\infty = (a, a, \ldots)$.

A zatem, w terminach strategii i polityk, decyzja rozmyta (21) przyjmuje postać

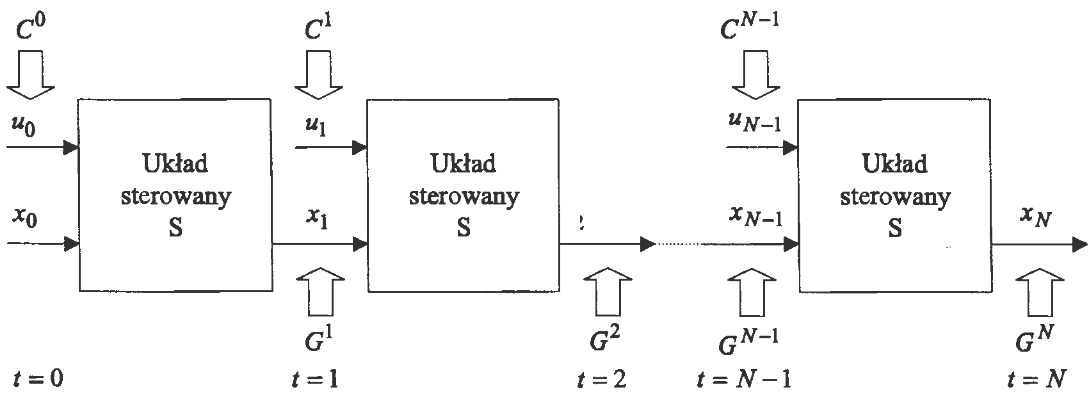
$$\mu_D(A \mid x_0) = \mu_{C^0}(a_0(x_0)) \wedge \cdots \cdots \wedge \mu_{C^{N-1}}(a_{N-1}(x_{N-1})) \wedge \mu_{G^N}(x_N) \quad (24)$$

a zadanie sprowadza się do znalezienia *strategii optymalnej* $A^* = (a_0^*, \ldots, a_{N-1}^*)$, takiej że

$$\mu_D(A^* \mid x_0) = \max_A \mu_D(A \mid x_0) \quad (25)$$

z uporządkowaniem między strategiami $A' \succeq A'' \Longleftrightarrow \mu_D(A' \mid x_0) \geq \mu_D(A'' \mid x_0)$, dla każdego $x_0 \in X$. *Strategia optymalna* A^* spełnia to oczywiście dla każdej A, tzn. $A^* \succeq A, \forall A$.

Możliwe są tu liczne rozszerzenia, głównie dla różnych: (1) czasów zakończenia, (2) układów sterowanych i (3) typów decyzji rozmytej (operacji agregacji). Tu omówimy przypadek sterowania układem stochastycznym przy skończonym, ustalonym z góry i nieskończonym czasie zakończenia procesu.



Rysunek 1. Istota wieloetapowego sterowania w rozmytym otoczeniu (w warunkach rozmytości)

4. Wieloetapowe sterowania rozmyte układem deterministycznym z ustalonym z góry, skończonym czasem zakończenia

Przypadek z układem deterministycznym ma znaczenie podstawowe jako punkt wyjścia. Dynamika deterministycznego układu sterowanego jest dana przez $x_{t+1} = f(x_t, u_t)$, $t = 0, 1, \ldots$, $x_t, x_{t+1} \in X = \{s_1, \ldots, s_n\}$, $u_t \in U = \{c_1, \ldots, c_m\}$. Stanem początkowym jest $x_0 \in X$, a czas zakończenia $N < \infty$ jest skończony, znany i zadany z góry. Na każdym t, na $u_t \in U$ nałożone jest ograniczenie rozmyte, $\mu_{C^t}(u_t)$, a na stan końcowy $x_N \in X$, nałożony jest cel rozmyty, $\mu_{G^N}(x_N)$.

Decyzja rozmyte ma postać

$$\mu_D(u_0, \ldots, u_{N-1} \mid x_0) = \mu_{C^0}(u_0) \wedge \ldots \ldots \wedge \mu_{C^{N-1}}(u_{N-1}) \wedge \mu_{G^N}(x_N) \tag{26}$$

gdzie $x_N = f(x_{N-1}, u_{N-1}) = f[f(x_{N-2}, u_{N-2}), u_0]$ itd.

Zadanie polega na znalezieniu *optymalnego ciągu sterowań*, $u_0^*, \ldots, u_{N-1}^*$; $u_t^* \in U, t = 0, \ldots, N-1$, takiego że

$$\begin{aligned}\mu_D(u_0^*, \ldots, u_{N-1}^* \mid x_0) &= \\ &= \max_{u_0, \ldots, u_{N-1}} \mu_D(u_0, \ldots, u_{N-1} \mid x_0) = \\ &= \max_{u_0, \ldots, u_{N-1}} [\mu_{C^0}(u_0) \wedge \ldots \\ &\ldots \wedge \mu_{C^{N-1}}(u_{N-1}) \wedge \mu_{G^N}(x_N)]\end{aligned} \tag{27}$$

i można je rozwiązać z użyciem dwu tradycyjnych technik: programowania dynamicznego i metody podziału i oszacowań oraz dwu nowych podejść: algorytmu genetycznego i sieci neuronowej (por. Kacprzyk [22], [23]). Poniżej wspomnimy tylko zastosowanie programowania dynamicznego zaproponowane u Bellmana i Zadeha [2].

Przepiszmy zadanie jako: znaleźć $u_0^*, \ldots, u_{N-1}^*$, taki że

$$\begin{aligned}\mu_D(u_0^*, \ldots, u_{N-1} \mid x_0) &= \\ &= \max_{u_0, \ldots, u_{N-1}} [\mu_{C^0}(u_0) \wedge \ldots \wedge \mu_{C^{N-1}}(u_{N-1}) \wedge \\ &\wedge \mu_{G^N}(f(x_{N-1}, u_{N-1}))]\end{aligned} \tag{28}$$

Struktura (28) umożliwia zastosowanie programowania dynamicznego, a układ równań rekurencyjnych programowania dynamicznego to:

$$\begin{cases} \mu_{G^{N-i}}(x_{N-i}) = \max_{u_{N-i}} [\mu_{C^{N-i}}(u_{N-i}) \wedge \\ \quad \wedge \mu_{G^{N-i+1}}(x_{N-i+1})] \\ x_{N-i+1} = \mathrm{f}(x_{N-i}, u_{N-i}); i = 0, 1, \ldots, N \end{cases} \tag{29}$$

gdzie $\mu_{G^{N-i}}(x_{N-i})$ to cel rozmyty na $t = N - i$ indukowany przez cel rozmyty na $t = N - i + 1, i = 0, 1, \ldots, N$.

Poszukiwane $u_0^*, \ldots, u_{N-1}^*$ są dane przez kolejne maksymalne wartości u_{N-i}, $i = 1, \ldots, N$ w (29), a każdą u_{N-i}^* otrzymuje się oczywiście jako funkcję stanu x_{N-i}, czyli otrzymujemy (*politykę optymalną*), $a_{N-i}(.)$, $u_{N-i}^* = a_{n-i}^*(x_{N-i})$, $i = 1, \ldots, N$. Rozwiązanie optymalne, istnieje, jeżeli jest przynajmniej jeden $u_0, \ldots, u_{N-1}$, taki że $\mu_D(u_0, \ldots, u_{N-1} \mid x_0) > 0$.

5. Wieloetapowe sterowanie rozmyte układem stochastycznym z ustalonym z góry, skończonym czasem zakończenia

Stochastyczny układ sterowany będzie tu łańcucem Markowa o przejściach stanów określonych przez $p(x_{t+1} \mid x_t, u_t)$; $t = 0, 1, \ldots$.

Stosuje się dwa następujące podstawowe sformułowania zadania, które nie są oczywiście równoważne:

- sformułowanie Bellmana i Zadeha [2]: znaleźć $u_0^*, \ldots, u_{N-1}^*$ maksymalizujący prawdopodobieństwo osiągnięcia celu rozmytego pod warunkiem spełnienia ograniczeń rozmytych, tzn.

$$\begin{aligned}\mu_D(u_0^*, \ldots, u_{N-1}^* \mid x_0) &= \\ &= \max_{u_0, \ldots, u_{N-1}} [\mu_{C^0}(u_0) \wedge \ldots \\ &\ldots \wedge \mu_{C^{N-1}}(u_{N-1}) \wedge E\mu_{G^N}(x_N)]\end{aligned} \tag{30}$$

- sformułowanie Kacprzyka i Staniewskiego [28] (to sformułowanie było przyjęte i rozwinięte w większości nowoczesnych podejść, np. u Iwamoto [13], Iwamoto,Tsurusaki i Fujita [15], Iwamoto i Sniedovicha [14] Yoshidy [33], [34] itd.): znaleźć $u_0^*, \ldots, u_{N-1}^*$ maksymalizujący wartość oczekiwaną decyzji rozmytej, tzn.

$$\begin{aligned}&\mu_D(u_0^*,\ldots,u_{N-1}^*\mid x_0)=\\&=\max_{u_0,\ldots,u_{N-1}} E\mu_D(u_0,\ldots,u_{N-1}\mid x_0)=\\&=\max_{u_0,\ldots,u_{N-1}} E[\mu_{C^0}(u_0)\wedge\ldots\\&\ldots\wedge\mu_{C^{N-1}}(u_{N-1})\wedge\mu_{G^N}(x_N)]\end{aligned}\tag{31}$$

W obu sformułowaniach szukamy strategii optymalnej $A^*=(a_0^*,\ldots,a_{N-1}^*)$, gdzie – jeśli tylko możliwe – $a_t^*: X\longrightarrow U$, takie że $u_t^*=a_0^*(x_t)$, $t=0,1,\ldots,N-1$, jest polityką optymalną na t.

5.1. Sformułowanie Bellmana i Zadeha

Struktura zadania jest taka sama jak (27), dla deterministycznego układu sterowanego. Otrzymujemy następujący układ równań rekurencyjnych programowania dynamicznego:

$$\begin{cases}\mu_{G^{N-i}}(x_{N-i})=\max_{u_{N-i}}[\mu_{C^{N-i}}(u_{N-i})\wedge\\\qquad\wedge E\mu_{G^{N-i+1}}(x_{N-i+1})]\\E\mu_{G^{N-i+1}}(x_{N-i+1})=\\\qquad=\sum_{x_{N-i}\in X}p(x_{N-i+1}\mid x_{N-i},u_{N-i})\times\\\qquad\times\mu_{G^{N-i+1}}(x_{N-i+1})\\i=1,\ldots,N\end{cases}\tag{32}$$

a kolejne maksymalizujące wartości u_{N-i}, u_{N-i}^*, $i=1,2,\ldots,N$, określają poszukiwany optymalny ciąg sterowań. Otrzymujemy znów optymalne polityki a_{N-i}^*, takie że $u_{N-i}^*=a_{N-i}^*(x_{N-i})$, $i=1,\ldots,N$. Są one markowskie, tzn. wyrażają aktualne sterowanie tylko w funkcji aktualnego stanu.

Rozszerzenie powyższego modelu na rozmyte prawdopodobieństwo zdarzenia rozmytego zaproponowano w pracach Kacprzyka [16], [19].

5.2. Sformułowanie Kacprzyka i Staniewskiego

Powyższe sformułowanie zadania jest nieco "dziwne", bo przecież bardziej naturalne byłoby szukać ciagu sterowań optymalnych maksymalizujących wartość oczekiwaną funkcji przynależności decyzji rozmytej. Niestety, od dawna wiadomo, że dla przypadku funkcji jakości (decyzji rozmytej) innej niż addytywna (ewentualnie średnia ważona) nie można otrzymać optymalnej polityki markowskiej (por. Howard [12], Iwamoto i Sniedovich [14]) i o tym na pewno wiedzieli Bellman i Zadeh. Takie bardziej "naturalne" sformułowanie przedstawimy poniżej.

Prawdopodobieństwa osiągnięcia x_N z x_0 poprzez $u_0,\ldots,u_{N-1}$ jest równe $p(x_1\mid x_0,u_0)\cdot p(x_2\mid x_1,u_1)\cdot\ldots\cdot p(x_N\mid x_{N-1},u_{N-1})$, a wartość oczekiwana zmiennej losowej $\mu_D(u_0,\ldots,u_{N-1}\mid x_0)$ jest równa

$$\begin{aligned}&E\mu_D(u_0,\ldots,u_{N-1}\mid x_0)=\\&=\sum_{(x_0,\ldots,x_N)\in X^{N+1}}\mu_D(u_0,\ldots,u_{N-1}\mid x_0)\times\\&\times[p(x_1\mid x_0,u_0)\cdot p(x_2\mid x_1,u_1)\cdot\ldots\\&\ldots\cdot p(x_N\mid x_{N-1},u_{N-1})]\end{aligned}\tag{33}$$

Wprowadźmy teraz ciąg funkcji $\hat{h}_i,h_j$, $i=0,1,\ldots,N$, $j=1,\ldots,N-1$:

$$\begin{cases}\hat{h}_N(x_N,u_0,\ldots,u_{N-1})=\mu_D(u_0,\ldots,u_{N-1}\mid x_0)\\\cdots\cdots\cdots\\\hat{h}_k(x_k,u_0,\ldots,u_{k-1})=\max_{u_k}h_k(x_k,u_0,\ldots,u_k)\\h_{k-1}(x_{k-1},u_0,\ldots,u_{k-1})=\\=\sum_{x_k\in X}\hat{h}_k(x_k,u_0,\ldots,u_{k-1})\cdot p(x_k\mid x_{k-1},u_{k-1})\\\cdots\cdots\cdots\\\hat{h}_0(x_0)=\max_{u_0}h_0(x_0,u_0)\end{cases}\tag{34}$$

czyli $h_{k-1}(.)$ jest wartością oczekiwaną $E\mu_D(.\mid x_0)$ przy założeniu optymalnych sterowań od $t=k$ do końca ($t=N$), $u_k^*,\ldots,u_{N-1}^*$. Wtedy $\hat{h}_k$ oznacza najwyższą oczekiwaną wartość decyzji rozmytej, jaką można osiągnąć, jeżeli jesteśmy na bieżącym t.

Można pokazać, że ponieważ X i U są skończone, to istnieją funkcje $w_k: X\times\underbrace{U\times\cdots\times U}_{k}\longrightarrow U$, $k=0,1,\ldots,N-1$, takie że $\hat{h}_k(x_k,u_0,\ldots,u_{k-1})=h_k[x_k,u_0,\ldots,u_{k-1},w_k(x_k,u_0,\ldots,u_{k-1})]$.

Jeżeli wprowadzimy ciąg funkcji $g_k^*:\underbrace{X\times\cdots\times X}_{k}\longrightarrow U$, $k=0,1,\ldots,N-1$, taki że

$$\begin{aligned}&g_k^*(x_0,\ldots,x_k)=\\&=w_k[x_k,g_0^*(x_0),\ldots,g_{k-1}^*(x_0,\ldots,x_{k-1})]\end{aligned}\tag{35}$$

gdzie $g_0^*(x_0)=w_0(x_0)$, to (Kacprzyk [22],[23]): g_k^* są poszukiwanymi politykami optymalnymi, tzn. $u_k^*=g_k^*(x_0,\ldots,x_k)$, $k=0,\ldots,N-1$, a ponadto $h_0(x_0)=\max_{u_0,\ldots,u_{N-1}}E\mu_D(u_0,\ldots,u_{N-1}\mid x_0)$, gdzie polityka odnosi sterowanie optymalne nie tylko do aktualnego stanu, ale także do "podsumowania" dotychczasowej trajektorii, jak zobaczymy dalej.

Te własności prowadzą do algorytmu [Kacprzyk [22], [23]] określania strategii optymalnej, który sprowadza się do kolejnego stosowania układu równań (34) dla sterowań reprezentowanych przez polityki optymalne (35). Oczywiście, otrzymane polityki optymalne nie są markowskie! Później Iwamoto [13] oraz Iwamoto i Sniedovich [14] wrócili do powyższego sformułowania i podali pogłębioną analizę algorytmu wyznaczania optymalnych polityk niemarkowskich oraz ulepszony algorytm.

6. Wieloetapowe sterowanie rozmyte układem stochastycznym z nieskończonym czasem zakończenia

Dotychczas czas zakończenia procesu (horyzont) był skończony, $t=0,1,\ldots,N$, a rozwiązywanie wymagało iteracji po kolejnych t, od 0 do N. Miało to oczywiście sens, gdy liczba etapów sterowania była niezbyt duża, czyli gdy proces sterowania nie był zbyt długi, a przede wszystkim, gdy zmienność w czasie była niewielka. Istnieje jednak wiele procesów sterowania, które zachodzą na bardzo długim horyzoncie, a przy tym mają na celu jedynie utrzymanie np. aktualnego poziomu działalności, jak np. w wielu problemach ekonomicznych, społecznych itp., gdy np. nie można oczekiwać rozwoju, a co najwyżej stagnacji. Wtedy należałoby raczej założyć *nieskończony czas zakończenia* i poszukiwać *optymalnej strategii stacjonarnej*. Oczywiście, iteracje po poszczególnych etapach należy zastąpić schematem iteracyjnych o skończonej liczbie kroków [por. np. książki Howarda [12], Bertsekasa [3], Bertsekasa i Shreve'a [4]], poprzez tzw. *metodę iteracji polityk*. Rozwiązanie, nawet dla horyzontu nieskończonego, otrzymuje się w skończonej liczbie iteracji.

W przypadku wieloetapowego podejmowania decyzji (sterowania) w warunkach rozmytości, zadanie z nieskończonym czasem zakończenia zostało po raz pierwszy sformułowane i rozwiązane przez Kacprzyka, Staniewskiego i Safteruka [27], a następnie rozszerzone np. w książkach Kacprzyka [22], [23]. Pokazano tam, że można zastosować pewien schemat interacji polityk. Ostatnio prace w tym zakresie prowadzi głównie Yoshida [33], [34], przede wszystkim dla zadań planowania finansowego.

Na każdym t, $t = 0, 1 \ldots$, na $u_t \in U = \{c_1, \ldots, c_m\}$ jest nałożone *zależne od stanu ograniczenie rozmyte* $\mu_C(u_t \mid x_t)$. Cel rozmyty $\mu_G(x_t)$, jest taki sam dla wszystkich t. Decyzją rozmytą jest

$$\mu_D(u_0, u_1, \ldots \mid x_0) = = \lim_{N\to\infty} \bigwedge_{t=0}^{N} [\mu_C(u_t \mid x_t) \wedge \mu_G(x_{t+1})] \quad (36)$$

lub, w przypadku użycia dyskonta, $b > 1$:

$$\mu_D(u_0, u_1, \ldots \mid x_0, b) = = \lim_{N\to\infty} \bigwedge_{t=0}^{N} b^t[\mu_C(u_t \mid x_t) \wedge \mu_G(x_{t+1})] \quad (37)$$

Zadanie polega na znalezieniu optymalnej strategii stacjonarnej $a^*_\infty = (a^*, a^*, \ldots)$, takiej że

$$\mu_D(a^*_\infty \mid x_0) = = \max_{a_\infty} \lim_{N\to\infty} \bigwedge_{t=0}^{N} [\mu_C(a(x_t) \mid x_t) \wedge \mu_G(x_{t+1})] \quad (38)$$

a w przypadku dyskonta ($b > 1$):

$$\mu_D(a^*_\infty \mid x_0, b) = = \max_{a_\infty} \lim_{N\to\infty} \bigwedge_{t=0}^{N} b^t[\mu_C(a(x_t) \mid x_t) \wedge \wedge \mu_G(x_{t+1})] \quad (39)$$

z uporządkowaniem w zbiorze strategii: $a'_\infty \succeq a''_\infty \Longleftrightarrow \mu_D(a_\infty{'} \mid x_0) \geq \mu_D(a_\infty{''} \mid x_0)$, dla każdego $x_0 \in X$, a $a^*_\infty \succeq a_\infty$, dla każdej a_∞.

Dla uproszczenia zakładamy tu, że decyzja rozmyta jest bez dyskonta. Poszukujemy $a^*_\infty = (a^*, a^*, \ldots)$, takiej że

$$E\mu_D(a^*_\infty \mid x_0) = \max_{a_\infty} E\mu_D(a_\infty \mid x_0) \quad (40)$$

z naturalnym uporządkowaniem w zbiorze strategii $a'_\infty \succeq a''_\infty \Longleftrightarrow E\,\mu_D(a'_\infty \mid x_0) \geq E\,\mu_D(a'' \mid x_0)$, dla każdego $x_0 \in X$ oraz, oczywiście, $a^*_\infty \succeq a_\infty$, dla każdej a_∞. Przyjmujemy tu więc sformułowanie Kacprzyka i Staniewskiego, bo podejście Bellmana i Zadeha nie ma oczywiście sensu w przypadku nieskończonego horyzontu.

Nietrudno zauważyć, że proces rozwiązywania obejmuje dwie fazy: (1) wyznaczenie $E\,\mu_D(a_\infty \mid x_0)$ oraz (2) wyznaczenie a^*_∞. Wyznaczenie $E\,\mu_D(a_\infty \mid x_0)$ dla decyzji rozmytej typu minimum nie jest zagadnieniem trywialnym – szczegóły można znaleźć w książkach Kacprzyka [22], [23].

Algorytm wyznaczanie optymalnej strategii stacjonarnej pierwsi zaproponowali Kacprzyk, Safteruk i Staniewski [27], a jego rozszerzenia i analizę można też znaleźć w książkach Kacprzyka [17], [18], [22], [23]. Przedstawimy teraz główne elementy analizy i rozwiązywania tego zadania. Ulepszone algorytmy zaproponował potem Yoshida (por. [33] [34]).

Oznaczmy najpierw h_t=(x_0, u_0, x_1, u_1, ..., x_{t-1}, u_{t-1}, x_t), t=0,1,..., gdzie $h_0 = (x_0)$, a $h = (x_0, u_0, x_1, u_1, x_2, \ldots)$. A więc, h_t reprezentuje trajektorię od $t = 0$ do t, a h reprezentuje "całkowitą" trajektorię. Oznaczmy przez $H_t = \{h_t\}$, dla każdego t, zbiór realizacji trajektorii h_t. Politykę określamy jako $a_t : H_t \longrightarrow U$, t=0, 1, ..., taką że $u_t = a_t(h_t) = a_t(x_0, u_0, x_1, \ldots, x_{t-1}, u_{t-1}, x_t)$, t=0, 1, ..., gdzie u_0=$a_0(h_0) = a_0(x_0)$.

Dla uproszczenia oznaczmy $r_{t-1} = r(x_{t-1}, a_{t-1}(h_{t-1}), x_t) = \mu_C(a_{t-1}(h_{t-1}) \mid x_{t-1}) \wedge \mu_G(x_t)$ oraz $R(h_t) = r_0 \wedge \ldots \wedge r_{t-1}$ i $R(h) = r_0 \wedge r_1 \wedge \ldots = \lim_{t\longrightarrow\infty} R(h_t)$.

Oznaczmy przez $v^T(h_t, A)$ wartość oczekiwaną $R(h_t)$, dla pewnego $T < \infty$, dla pewnych h_t i $A = (a_0, a_1, \ldots, a_t)$, $t < T$. Jest to oczywiście równe

$$\begin{aligned} v^T(h_t, A) &= \\ &= \sum_{(x_{t+1}, x_{t+2} \ldots, x_T)} [(R(h_t) \wedge r_t \wedge \ldots \wedge r_{T-1}) \cdot \\ &\quad \cdot p(x_{t+1} \mid x_t, a(h_t)) \times \cdots \\ &\quad \cdots \times p(x_T \mid x_{T-1}, a_{T-1}(h_{T-1}))] = \\ &= \sum_{(x_{t+1}, x_{t+2}, \ldots, x_T)} [(R(h_t) \wedge \bigwedge_{k=t}^{T-1} r_k) \times \\ &\quad \times \prod_{q=t}^{T-1} p(x_{q+1} \mid x_q, a_q(h_q))] \end{aligned} \quad (41)$$

Ponadto oznaczmy: $v(h_T, A) = \lim_{T\longrightarrow\infty} v^T(h_T, A)$, $f^T(h_t) = \max_A v^T(h_t, A)$ i $f(h_t) = \lim_{T\longrightarrow\infty} f^T(h_t)$, przy czym maksymalizacja po A odbywa się tylko po zbiorze polityk $\{a_t, a_{t+1}, \ldots, a_{T-1}\}$, ponieważ polityki a_0, a_1, ..., a_{t-1} są ustalone, bo h_t jest z założenia dane. Ponieważ, co wynika z zastosowania "$\wedge$", ciągi $(v^T(h_t, A))_{t=0,1,\ldots}$ i $(f^T(h_t))_{t=0,1,\ldots}$ są nierosnące, więc te granice istnieją.

Strategię A^* nazywamy *strategią optymalną*, jeżeli $v(h_t, A^*) = f(h_t)$, $t = 0, 1, \ldots$.

Potrzebujemy teraz *równania funkcyjnego*, które tu ma postać, dla każdego t = 0, 1, ..., h_t i A:

$$v(h_t, A) = \sum_{x_{t+1}} v(x_{t+1}, A) \cdot p(x_{t+1} \mid x_t, a_t(h_t)) \quad (42)$$

Mamy teraz *warunek optymalności*: dla każdego t i h_t, zachodzi

$$f(h_t) = \max_{a_t} \sum_{x_{t+1}} f(h_{t+1}) \cdot p(x_{t+1} \mid x_t, a_t(h_t)) \quad (43)$$

$\hat{A} = (\hat{a}_0, \hat{a}_1, \ldots)$ nazywamy *strategią zachowawczą*, jeżeli $f(h_t) = \sum_{x_{t+1}} f(h_{t+1}) \cdot p(x_{t+1} \mid x_t, \hat{a}_t(h_t))$. Można pokazać, że: (1) zawsze istnieje strategia zachowawcza, (2) każda strategia zachowawcza jest optymalna, czyli istnieje strategia optymalna.

Strategię $\breve{A} = (\breve{a}_0, \breve{a}_1, \ldots)$ nazywamy *niepoprawialną* (w jednym kroku), jeżeli

$$v(h_t, \breve{A}) = \max_{A} \sum_{x_{t+1}} v(h_{t+1}, \breve{A}) \cdot p(x_{t+1} \mid x_t, \breve{a}_t(h_t)) \quad (44)$$

czyli taką, która spełnia warunek optymalności (43). Każda strategia optymalna jest niepoprawialna, ale nie odwrotnie.

Jeżeli $A = (a_0, a_1, \ldots)$ nie jest niepoprawialna, to można skonstruować strategię $\overline{A} = (\overline{a}_0, \overline{a}_1, \ldots)$, taką że $v(h_t, A) \leq \sum_{x_{t+1}} v(h_{t+1}, A) \cdot p(x_{t+1} \mid x_t, \overline{a}_t(h_t))$, o której mówimy, że *poprawia* A (w jednym kroku).

Mamy więc, że jeżeli strategia $\overline{A}$ poprawia strategię A, to $v(h_t, A) \leq v(h_t, \overline{A})$, dla każdego t i h_t, z "$<$" przynajmniej dla jednego t i h_t.

Zatem, jeżeli można skonstruować procedurę iteracyjną znajdującą kolejne strategie $A_0, A_1, \ldots$, takie że A_{i+1} poprawia A_i, dla $i = 0, 1, \ldots$, to powinna się ona zakończyć, kiedy zostanie znaleziona strategia niepoprawialna. Jest ona na pewno bardzo dobrym rozwiązaniem, chociaż nie musi być niestety strategią optymalną.

Interesują nas strategie stacjonarne, więc rozpatrzymy pojęcie *stacjonarności*. Oznaczmy najpierw $w(h_t) = (x_t, R(h_t))$, gdzie $R(h_t)$ nazwiemy *podsumowaniem* trajektorii h_t. Zarówno x_t jak $R(h_t)$ mogą przyjmować skończoną liczbę wartości; $w(h_t)$ może więc także przyjmować skończoną liczbę wartości (par wartości z $X \times [0, 1]$). Załóżmy, że $w(h_t) \in W = \{y_1, \ldots, y_q\}$, gdzie W jest zbiorem (różnych) wartości $w(h_t)$.

Strategię $A = (a_0, a_1, \ldots)$ nazywamy *stacjonarną*, $a_\infty = (a, a, \ldots)$, jeżeli $a_t(h_t) = a_{t'}(h_{t'})$, dla każdych h_t i $h_{t'}$, takich że $w(h_t) = w(h_{t'})$.

Mamy więc ważną właściwość polityk stacjonarnych. Dla każdej strategii stacjonarnej a_∞, jeżeli $w(h_t) = w(h_{t'})$, to $v(h_t, a_\infty) = v(h_{t'}, a_\infty)$ i $f(h_t) = f(h_{t'})$.

A więc, istnieje stacjonarna strategia zachowawcza, istnieje optymalna strategia stacjonarna, poprawa strategii (w jednym kroku) daje coraz lepsze strategie, a optymalnej strategii stacjonarnej należy poszukiwać wśród równoważnych strategii otrzymanych w wyniku poprawy strategii, czyli wśród strategii niepoprawialnych.

Te wyniki prowadzą do (sub)optymalnego algorytmu typu iteracji polityk do rozwiązywania rozpatrywanego zadania. Ten algorytm daje w wyniku strategie niepoprawialne typu (44). Strategia optymalna jest wśród nich. Jeżeli ich liczba nie jest zbyt wysoka, to można dokonoć pełnego przeglądu i znaleźć optymalną, a jeżeli nie, to należy raczej zastosować jakąś metodę heurystyczną niepełnego przeglądu. Podobne wyniki otrzymał Yoshida (por. np. [33], [34]. Ogólnie, w tych późniejszych pracach udało się zmniejszyć zbiór strategii niepoprawialnych.

7. Przykład zastosowania

Wieloetapowe sterowanie rozmyte znalazło wiele zastosowań, m.in. do: społeczno-ekonomicznego rozwoju regionalnego, zapobiegania powodziom, planowania prac badawczo-rozwojowych (R&D), wyznaczania harmonogramu włączania jednostek generujących w systemie energetycznym, sterowania dozowaniem środków anestezjologicznych podczas operacji, przydziału zasobów, sterowania zapasami itp. co omówiono w książkach Kacprzyka [22], [23].

Rozpatrzymy tu krótko model zrównoważonego rozwoju regionalnego, zaproponowany w pracach Kacprzyka i Straszaka [29], a potem Kacprzyka [24]. Jest to typowe "miękkie" zadanie analizy systemowej [por. Checkland [7]].

Rozważamy region rolniczy, w którym występuje migracja młodych mieszkańców do pobliskich ośrodków miejskich, co prowadzi do starzenia się populacji i w rezulacie do upadku społeczno-ekonomicznego regionu. Ta migracja jest spowowana przede wszystkim przez niską percepcję *jakości życia*. Aby więc zatrzymać ten upadek, należy podwyższyć (znacznie) jakość życia poprzez pewne nakłady zewnętrzne. Zadanie polega na znalezieniu wielkości tych inwestycji, przy spełnieniu nałożonych na nie ograniczeń, oraz ich rozkładu w czasie (na określonym horyzoncie planowania), tak aby najlepiej osiągnąć cele rozwoju i ich percepcję społeczną.

Region reprezentuje się jako społeczno-ekonomiczny dynamiczny układ sterowany, którego stan na etapie $t-1$, X_{t-1}, jest scharakteryzowany przez zbiór istotnych społeczno-ekonomicznych wskaźników jakości życia. Tak więc, sterowanie (nakład) na etapie $t-1$, u_{t-1}, zmienia stan (wartość tych wskaźników) na X_t. Powtarza się to dla poszczególnych etapów $t = 1, \ldots, N$, gdzie N jest zakładanym horyzontem planowania. Oczywiście, ze względu na wielki stopień niepewności te przejścia stanów mają charakter stochastyczny, a więc mamy stochastyczny układ sterowany.

Oceny "dobroci" etapu rozwoju t dokonuje się biorąc pod uwagę zarówno "dobroć" przyłożonego sterowania u_{t-1}, jak też "dobroć" osiągniętego stanu X_t; ta pierwsza ocena dotyczy tego, jak dobrze są spełnione pewne ograniczenia, natomiast ta druga tego, jak dobrze są osiągnięte pewne cele. Ocenę tę powtarza się dla $t = 1, \ldots, N$. A z innej perspektywy, taka ocena dotyczy zarówno "kosztów", czyli u_{t-1}, jak "efektów", czyli X_t. Zadanie sterowania polega na znalezieniu takiego ciągu inwestycji (sterowań), dla którego podana powyżej ocena rozwoju jest najbardziej pozytywna.

Problem polega na tym, jak właściwie sformułować i ocenić te koszty i efekty. Po pierwsze, przedstawmy system społeczno-ekonomiczny jako dynamiczny układ sterowany. Jego stan (wyjście) X_t utożsamia się ze *wskaźnikiem jakości życia*, który składa się z siedmiu *indykato-*

rów jakości życia: x_t^1 – jakość ekonomiczna (np. zarobki, dochody, ...), x_t^2 – jakość środowiskowa, x_t^3 – jakość mieszkań, x_t^4 – jakość służby zdrowia, x_t^5 – jakość infrastruktury, x_t^6 – możliwość znalezienia zatrudnienia, x_t^7 – możliwość czasu wolnego, czyli $X_t = [x_t^1, \ldots, x_t^7]$.

Sterowanie na $t-1$, u_{t-1}, jest inwestycją mającą na celu rozwój regionu. Na sterowanie u_{t-1} jest nałożone ograniczenie rozmyte $\mu_{C^{t-1}}(u_{t-1})$. Sterowanie (nakłady) na $t-1$, u_{t-1}, dzieli się na siedem składników, $u_{t-1}^1, \ldots, u_{t-1}^7$, których celem jest poprawienie wartości $x_{t-1}^1, \ldots, x_{t-1}^7$. Dynamikę opisuje $x_t^i = f_{t-1}^i(x_{t-1}^i, u_{t-1}^i)$, $i = 1, \ldots, 7$; $t = 1, \ldots, N$, które określa przejście z X_{t-1} do X_t przy inwestycjach $u_{t-1}^1, \ldots, u_{t-1}^7$. Takie równanie przejść stanów można wyznaczyć np. używając opinii ekspertów, danych z przeszłości itd. Oczywiście, właściwie mamy $p(x_t^i \mid x_{t-1}, u_{t-1})$.

Należy teraz ocenić dobroć przejścia z X_{t-1} do X_t. Istotą jego oceny jest więc dostarczenie pewnych miar tego, jak dobrze są spełnione ustalone cele. Powinno to być oczywiście odniesione do wyłożonych funduszy. Ponadto, jak wskazuje praktyka, powyższa ocena "kosztów" i "efektów" może często nie odpowiadać społecznemu odczuciu "dobroci" rozwoju. Te dwa ważne aspekty to: (1) efektywnością rozwoju i (2) stabilnością rozwoju.

Efektywność rozwoju regionalnego oznacza, jak dobrze są spełnione ograniczenia rozmyte i jak dobrze są osiągnięte cele rozmyte. Na to składają się dwa aspekty: (a) efektywność danego etapu rozwoju i (b) efektywność całej trajektorii (rozwoju), czyli wszystkich etapów na całym założonym horyzoncie plamowania. Rozpatrzmy teraz pierwszy element, czyli *efektywność pojedynczego etapu rozwoju*. Ma ona zarówno aspekt *obiektywny* jak *subiektywny*. Obiektywna ocena etapu t sprowadza się do określenia, jak dobrze są spełnione ograniczenia rozmyte i jak dobrze są osiągnięte cele rozmyte.

Aby uzyskać obiektywną ocenę całego wskaźnika jakości życia na t, $X_t = [x_t^1, \ldots, x_t^7]$, należy zagregować oceny cząstkowe poszczególnych indykatorów jakości życia, Zauważmy na koniec, że ocena obiektywna z swej istocie dotyczy bardziej władz niż mieszkańców, ponieważ niejako "mechanicznie" sprawdza, jak wyglądają osiągnięte wartości indykatorów jakości życia w porównaniu z określonymi z góry pożądanymi poziomami. Ich percepcja przez mieszkańców może być jednak różna.

Ocena mieszkańców dobroci rozwoju dotyczy (percepcji) *zadowolenia społecznego* wynikającego z osiągniętego wskaźnika jakości życia. Jest ona oczywiście *subiektywna*. Osiągnięta wartość x_t^i implikuje jego częściowe *zadowolenie społeczne*, s_t^i, którego interpretacja jest w zasadzie taka sama, jak dla oceny obiektywnej. Zadowolenie społeczne na etapie rozwoju t jest więc agregacją poszczególnych częściowych zadowoleń. Na zadowolenie społeczne s_t jest teraz nałożony subiektywny cel rozmyty $\mu_{G_s^t}(s_t)$.

Przez efektywność etapu rozwoju t rozumie się zasadniczo relację tego, co zostało osiągnięte (w sensie wskaźników jakości życia i odpowiadających im zadowoleń społecznych) do tego, co za to zapłacono (w sensie inwestycji); jest to więc pewna relacja typu koszty–efekty. Jest to oczywiście wynikiem agregacji odzwierciedlającej istotę kompromisu między interesami władz (dla których ważne są ograniczenia i cele rozmyte) i mieszkańców (dla których ważne są subiektywne podcele rozmyte, a także do pewnego stopnia obiektywne cele rozmyte). Użycie operacji minimum jest odbiciem podejścia typu bezpiecznego, a więc "bardziej sprawiedliwego" kompromisu, tzn. do przyjęcia dla obu partii, podczas gdy użycie maksimum odzwierciedla postawę odwrotną; pozostałe operacje odpowiadają przypadkom pośrednim.

Na koniec miary efektywności poszczególnych etapów należy zagregować, aby otrzymać rozmytą miarę efektywności dla całego rozwoju (czyli na całym horyzoncie planowania N). W tym przypadku operacje agregacji odzwierciedlają opinię dotyczącą tego, jak ważne jest to, co zajdzie w przyszłości, od pesymistycznej i ostrożnej (dla "$\wedge$") do optymistycznej (dla "$\vee$"), przez wszystkie przypadki pośrednie. Analizę stosowanych tu operatorów agregacji, z punktu widzenia ich adekwatności i możliwości rozwiązania wynikających zadań zawarto w pracy Kacprzyka [24].

Mamy więc rozmytą miarę efektywności całej ścieżki rozwoju. Niestety, nawet jeżeli ocena rozwoju w sensie tej miary jest pozytywna, to jego aktualna percepcja – przez władze jak mieszkańców – może być znacznie mniej korzystna. Jedną z głównych przyczyn może być zbyt duża zmienność pewnych wskaźników rozwoju i jego charakterystyk. To ważne zagadnienie jest zwane stabilnością rozwoju regionalnego. Dotyczy ono zmienności (kluczowych) indykatorów, charakterystyk, warunków itp. rozwoju. Wiadomo z doświadczenia, a potwierdzają to badania psychologiczne, że ograniczona zmienność zwykle implikuje większą akceptowalność rozwoju, zarówno przez władze jak mieszkańców, niż wysoka zmienność. *Stabilność* w powyższym sensie obejmuje: (1) *stabilność trajektorii rozwoju*, która dotyczy zmienności wyników rozwoju, tzn. osiągniętych indykatorów jakości życia i ospowienich zadowoleń społecznych, oraz (2) *stabilność "polityki" rozwoju*, która obejmuje zmienność pewnych warunków wst'epnych rozwoju, czyli nałożonych ograniczeń i celów rozmytych oraz reguł podziału inwestycji. Jasne jest, że obie są pojęciami miękkimi i powinny być oceniane z użyciem narzędzi rozmytych. Ograniczenia na zmienność wprowadza sie poprez *limity zmienności*, podobne do ograniczeń rozmytych.

Praktycznie wszystkie wielkości występujące w powyższym modelu charakteryzują się wysokim stopniem niepewności. Uwzględnienie tego prowadzi do modelu wieloetapowego sterowania rozmytego ze stochastycznym układem sterowanym, a ponieważ trudno oczekiwać rozwoju w przypadku takiego regionu, więc wprowadzenie nieskończonego czasu zakończenia jest dobrym wyborem. Niestety, co też nietrudno zauważyć, prowadzi to do bardzo skomplikowanych, pojęciowo i obliczeniowo, modeli, bo np. każde z "prostych" równań przejść to przecież skomplikowany model np. demograficzny czy dot. ochrony "srodowiska".

Więcej szczegółów nt. tego modelu, innych modeli do rozwiązywania podobnych zagadnień, a także modeli dla wielu innych problemów można znaleźć u Kacprzyka [22], [23]. Ponadto, w pracy Kacprzyka [24] pokazano, że – na szczęście – prostsze operatory agregacji (głównie typu $\wedge$), dla których można zadanie rozwiązać, są uzasadnione w przypadku rozpatrywanego powyżej rozwoju regionalnego, a także innych podobnych zadań.

8. Uwagi końcowe

Praca zawiera przegląd podstawowych modeli wieloetapowego sterowania rozmytego przy założeniu, że układ sterowany jest stochastyczny (łańcuch Markowa). Podano też przykład zastosowania tego typu modeli do planowania regionalnego, a więc zadania intuicyjnie zrozumiałego.

Wśród możliwych dalszych kierunków rozwoju pokazanych modeli należy wymienić następujące. Po pierwsze, w ostatnim czasie pojawiło się wiele prac nt. określania prawdopodobieństw zdarzeń rozmytych jako liczb rozmytych, a nawet przedziałów rozmytych. Skomplikuje to niewątpliwie modele, ale pozwoli na badziej adekwatną reprezentację. Po drugie, wydaje się, że nowe perspektywy może otworzyć zastosowanie do opisu dynamiki układu sterowanego zmiennych niepewnych Bubnickiego [5], [6], co powinno umożliwić ujęcie w szerszy sposób niepewności przy opisie dynamiki.

Literatura

[1] Bellman R.E. (1957) *Dynamic Programming*. Princeton University Press.

[2] Bellman R.E., Zadeh L.A. (1970) Decision making in a fuzzy environment. *Management Science* 17: 141–164.

[3] Bertsekas D.P. (1976) *Dynamic Programming and Stochastic Control*. Academic Press, New York.

[4] Bertsekas D.P., Shreve S.E. (1978) *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, New York.

[5] Bubnicki Z. (2002) *Uncertain Logics, Variables and Systems*. Springer, Heidelberg and New York.

[6] Bubnicki Z. (2004) *Analysis and Decision Making in Uncertain Systems*. Springer, Heidelberg and New York.

[7] Checkland P.B. (1976) Science and the systems paradigm. *Int. J. of General Systems* 3: 127–134.

[8] Denardo E.V. (1982) *Dynamic Programming Models and Applications*. Prentice-Hall, Englewood Cliffs.

[9] Dompere K.K. (2004) *Cost-Benefit Analysis and the Theory of Fuzzy Decisions*. Ī. Fuzzy Value Theory, 2. *Identification and Measurement Theory*. Springer, Heidelberg and New York.

[10] Esogbue A.O., Fedrizzi M., Kacprzyk J. (1988) Fuzzy dynamic programming with stochastic systems. W J. Kacprzyk i M. Fedrizzi (pod red.): *Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making*. Springer–Verlag, Berlin and New York, str. 266–285.

[11] Esogbue A.O., Kacprzyk J. (1998) Fuzzy dynamic programming. W R. Słowiński, pod red.: *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, Kluwer, Boston, 281–307.

[12] Howard R.A. (1971) *Dynamic Probabilistic Systems*. Volumes 1 and 2, Wiley, New York.

[13] Iwamoto S. (2001) Fuzzy dynamic programming in stochastic environment, w Y. Yoshida, pod red.: *Dynamic Aspects in Fuzzy Decision Making*; Physica-Verlag, Heidelberg and New York, 27–51.

[14] Iwamoto S., Sniedovich M. (1998) Sequential decision making in fuzzy environment. *J. of Mathematical Analysis and Applications*. 222: 208–224.

[15] Iwamoto S., Tsurusaki K., Fujita T. (2001) On Markov policies for minimax decision processes. *J. of Mathematical Analysis and Applications*, 253: 58–78.

[16] Kacprzyk J. (1982) Control of a stochastic system in a fuzzy environment with Yager's probability of a fuzzy event. *Busefal* 12: 77–88.

[17] Kacprzyk J. (1983) *Wieloetapowe Podejmowanie Decyzji w Warunkach Rozmytości*. PWN, Warszawa–Łódź.

[18] Kacprzyk J. (1983) *Multistage Decision Making under Fuzziness*, Verlag TÜV Rheinland, Cologne.

[19] Kacprzyk J. (1984) Yager's probability of a fuzzy event in stochastic control under fuzziness. W M.M. Gupta i M. Sanchez (pod red.): *Fuzzy Information, Knowledge Representation and Decision Analysis*, Pergamon Press, Oxford, str. 379–384.

[20] Kacprzyk J. (1987) *Zbiory Rozmyte w Analizie Systemowej* PWN, Warszawa.

[21] Kacprzyk J. (1987) Stochastic systems in fuzzy environments: control. W M.G. Singh (pod red.): *Systems and Control Encyclopedia*, Pergamon Press, Oxford, str. 4657–4661.

[22] Kacprzyk J. (1997) *Multistage Fuzzy Control*, Wiley, Chichester.

[23] Kacprzyk J (2001) *Wieloetapowe Sterowanie Rozmyte*, WNT, Warszawa.

[24] Kacprzyk J. (2004) Fuzzy dynamic programming with stochastic systems under various aggregation operators: solvability and perceived meaning. W López-Díaz M., Gil M.A., Grzegorzewski P., Hryniewicz O., Lawry J. (pod red.): *Soft Methodology and Random Information Systems*. Springer, Heidelberg and New York, str. 551–558.

[25] Kacprzyk J., Esogbue A.O. (1996) Fuzzy dynamic programming: main developments and applications. *Fuzzy Sets and Systems.* 81: 31–45.

[26] Kacprzyk J., Fedrizzi M., pod red. (1988) *Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making*, Springer-Verlag, Berlin.

[27] Kacprzyk J., Safteruk K., Staniewski P. (1981) On the control of stochastic systems in a fuzzy environment over infinite horizon. *Systems Science* 7: 121–131.

[28] Kacprzyk J., Staniewski P. (1980) A new approach to the control of stochastic systems in a fuzzy environment. *Archiwum Automatyki i Telemechaniki* XXV: 433–443.

[29] Kacprzyk J., Straszak A. (1984) Determination of stable trajectories for integrated regional development using fuzzy decision models. *IEEE Trans. on Systems, Man and Cybernetics* SMC-14: 310–313.

[30] Puterman M.L. (1994) *Markov Decision Processes. Discrete Stochastic Dynamic Programming*. Wiley, New York.

[31] Sniedowich M. (1992) *Dynamic Programming.* Marcel Dekker, New York.

[32] Yager R.R. (1979) A note on probabilities of fuzzy events. *Information Sciences* 18: 113–129.

[33] Yoshida Y. (1994) Markov chains with a transition possibility measure and fuzzy dynamic programming. *Fuzzy Sets and Systems* 66: 39–57.

[34] Yoshida Y., pod red. (2001) *Dynamical Aspects in Fuzzy Decision Making*. Physica-Verlag, Heidelberg and New York.

[35] Zadeh L.A. (1965) Fuzzy sets. *Information and Control* 8: 338–353.

[36] Zadeh L.A. (1986) Fuzzy probabilities. *Information Processing and Management* 20: 363–372.

[37] Zadeh L.A. (1972) A rationale for fuzzy control. *Measurement and Control* 34: 3–4.

[38] Zadeh L.A., Kacprzyk J., pod red. (1992) *Fuzzy Logic for the Management of Uncertainty*, Wiley, New York.

[39] Zadeh L.A., Kacprzyk J., pod red. (1999) *Computing with Words in Information/Intelligent Systems. 1. Foundaiions, 2. Applications*. Physica-Verlag, Heidelberg and New York.

Instytut Badań Systemowych
Polskiej Akademii Nauk

ISBN 83-89475-02-2