INFORMATION PROCESSING LETTERS 2 (1973) 74–78. NORTH-HOLLAND PUBLISHING COMPANY

EXAMPLES OF FORMAL GRAMMARS WITH WEIGHTS

Masaharu MIZUMOTO, Junichi TOYODA and Kohkichi TANAKA
Dep. of Inf. and Comp. Sciences, Faculty of Engng. Science, Osaka University, Toyonaka, Osaka, Japan 560

Received March 1973
Revised version received June 1973

formal grammars with weights
weighted grammars
probabilistic grammars
context-free languages
fuzzy grammars
context sensitive languages

In [1] we have derived various kinds of formal grammars with weights by incorporating the algebra systems with formal grammar systems. There we evaluate the weights of sentences by making an element of the appropriate algebra (say, lattice ordered semigroup and distributive lattice) correspond to each rewriting rule of a formal grammar and by performing the operations of the algebras in the case of the application of the rewriting rules in a derivation. Among these sorts of formal grammars with weights are the well-known probabilistic grammars [2, 3], fuzzy grammars [4–7], and weighted grammars [2]. For example, probabilistic grammars are the formal grammars with weights such that the values in the range $[0, 1]$ are adopted as the weights (or probabilities) where, needless to say, the condition of probability is assumed to be satisfied, and + (addition) and × (multiplication) are used as the operations. In addition, weighted grammars can be formulated by extending the range $[0, 1]$ in probabilistic grammars to the set of non-negative real numbers $[0, \infty)$. A number of researchers are engaged in the studies on these grammars because these grammars have enhanced generative power.

In this paper, as a preliminary step, we shall give some interesting examples of formal grammars with weights in which the generative powers of languages can be enhanced by adopting other algebra systems such as boolean lattices and rings.

Definition. A *formal grammar with weights* is a system

$$G = (V_N, V_T, P, S, J, M, \mu) \tag{1}$$

where V_N, V_T, and S are defined as usual [8]. P is a set of productions such as

$$(r)\, u \to v\mu(r)\,. \tag{2}$$

$J = \{r\}$ is a set of labels. $u \to v$ is an ordinary rewriting rule. μ is a *weighting function* such as $\mu: J \to M$. M is a *weighting space*. The value $\mu(r)$, $r \in J$, represents the *weight* of the application of a rule r.

The expression

$$S \underset{r_1}{\overset{\mu(r_1)}{\longrightarrow}} \alpha_1 \underset{r_2}{\overset{\mu(r_2)}{\longrightarrow}} \alpha_2 \to \cdots \underset{r_m}{\overset{\mu(r_m)}{\longrightarrow}} x \tag{3}$$

will be referred to as a *weighted derivation chain* from S to x by the productions $r_1, r_2, \cdots, r_m$. In general, there is more than one weighted derivation chain from S to x.

Let V and $*$ be two operations† of weighting space M, then the weight $\mu_G(x)$ of the generation of x in V_T^* is defined as follows by using the weighted derivation chain of (3) and the operations V and $*$ (see ref. [1]).

† It should be noted that the operation V is the representative notation for a certain operation in a weighting space and therefore does not necessarily represent the operation of l.u.b. in lattices. The same holds for the operation $*$.

$$\mu_G(x) = V[\mu(r_1) * \mu(r_2) * \cdots * \mu(r_m)] \;, \tag{4}$$

where the operation V is taken over all the weighted derivation chains from S to x.

Let us define two kinds of languages with threshold λ ($\in M$) generated by a formal grammar G with weights as follows.

$$L(G, \lambda) = \{x \in \nabla_T^* \mid \mu_G(x) > \lambda\} \;, \tag{5}$$

$$L(G,=,\lambda) = \{x \in \nabla_T^* \mid \mu_G(x) = \lambda\} \;. \tag{6}$$

Now we shall give several interesting examples of formal grammars with weights whose generative power is enhanced by adopting an algebra such as boolean lattices and rings†.

First, let the weighting space M be the boolean lattice B.

Example 1. Consider the following formal grammar G with weights in which all the rewriting rules are of type 2 (or CF) form. $G = (V_N, V_T, P, S, J, M, \mu)$, where $V_N = \{S, A, B, C, D\}$, $V_T = \{a, b, c\}$, and $M = B$ (boolean lattice) as in fig. 1. Moreover let the operation V in (4) be l.u.b. in B and the operation $*$ be $\wedge$ (= g.l.b. in B). P is given as follows.

(1) $S \rightarrow AB$ $\quad x_1$,

(2) $S \rightarrow CD$ $\quad x_2$,

(3) $A \rightarrow aAb$ $\quad I$,

(4) $A \rightarrow ab$ $\quad I$,

(5) $B \rightarrow cB$ $\quad I$,

(6) $B \rightarrow c$ $\quad I$,

(7) $C \rightarrow aC$ $\quad I$,

(8) $C \rightarrow a$ $\quad I$,

(9) $D \rightarrow bDc$ $\quad I$,

(10) $D \rightarrow bc$ $\quad I$.

† There exist formal grammars with weights but without enhanced generative power. Among these are fuzzy grammars (see ref. [5]).

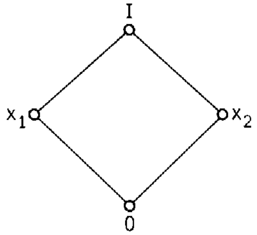

Fig. 1. Structure of M.

A string, say $a^2b^2c^2$ is obtained by the following weighted derivation chain.

$$\underline{S} \underset{1}{\overset{x_1}{\rightarrow}} \underline{A}B \underset{3}{\overset{I}{\rightarrow}} aAb\underline{B} \underset{5}{\overset{I}{\rightarrow}} a\underline{A}bcB \underset{4}{\overset{I}{\rightarrow}} a^2b^2c\underline{B} \underset{6}{\overset{I}{\rightarrow}} a^2b^2c^2 \;,$$

where the underbar in the intermediate strings represents the location where the next rule was applied. The weight of the generation of $a^2b^2c^2$ by this derivation is given as follows.

$$x_1 \wedge I \wedge I \wedge I \wedge I = x_1 \;.$$

Similarly, for the same string $a^2b^2c^2$, the following derivation is also possible.

$$\underline{S} \underset{2}{\overset{x_2}{\rightarrow}} \underline{C}D \underset{7}{\overset{I}{\rightarrow}} aC\underline{D} \underset{9}{\overset{I}{\rightarrow}} a\underline{C}bDc \underset{8}{\overset{I}{\rightarrow}} a^2b\underline{D}c \underset{10}{\overset{I}{\rightarrow}} a^2b^2c^2 \;.$$

In this case we have x_2. Furthermore we can also give the different weighted derivation chains of $a^2b^2c^2$, all of whose weights are shown to be x_1 or x_2.

Hence the weight of the generation of $a^2b^2c^2$ by this grammar G is given from (4) as follows.

$$\mu_G(a^2b^2c^2) = x_1 \vee x_2 = I \;.$$

Continuing in this manner we can obtain the weight $\mu_G(x)$ for each terminal string $x \in V_T^*$. Let $\{(x, \mu_G(x))\}$ be a set of ordered pairs of x ($\in V_T^*$) and its weight $\mu_G(x)$. Hence, for this grammar G, $\{(x, \mu_G(x))\}$ is given as follows.

$$\{(x, \mu_G(x))\} = \{(a^ib^ic^i, I) \mid i \geqslant 1\}$$
$$\cup \{(a^ib^ic^j, x_1) \mid i,j \geqslant 1, i \neq j\}$$
$$\cup \{(a^ib^jc^j, x_2) \mid i,j \geqslant 1, i \neq j\} \;.$$

Therefore we have from (5) and (6)

$$L(G,x_1) = L(G,x_2) = L(G,=,I) = \{a^i b^i c^i \mid i \geqslant 1\}\ . \quad (7)$$

It is noted that the languages generated by this grammar G on B are CS languages though G contains CF rules only.

Example 2. Assume that G is a formal grammar with weights which consists of type 3 rules as follows. Let $M = Z$ (set of integers) and let the operation V in (4) be replaced by $\times$ (multiplication) and $*$ by + (addition); M thus forms a ring. Productions are given by the following.

(1) $S \to aS$ $\quad +1$,

(2) $S \to aA$ $\quad +1$,

(3) $A \to bA$ $\quad -1$,

(4) $A \to b$ $\quad -1$,

A string, say, a^3b^2 can be generated by the following derivation chain.

$$S \underset{1}{\overset{+1}{\to}} aS \underset{1}{\overset{+1}{\to}} a^2S \underset{2}{\overset{+1}{\to}} a^3A \underset{3}{\overset{-1}{\to}} a^3bA \underset{4}{\overset{-1}{\to}} a^3b^2\ .$$

In this case the weight of a^3b^2 by this derivation chain is given as 1 (= + 1 + 1 + 1 − 1 − 1). As there exists no other derivation chain for a^3b^2, the weight $\mu_G(a^3b^2)$ of a^3b^2 by this grammar G is 1. In general we have

$$\{(x, \mu_G(x))\} = \{(a^i b^j, i-j) \mid i,j \geqslant 1\}\ .$$

Therefore the languages with threshold $\lambda = 0\ (\in Z)$ can be represented by (5) and (6) as follows.

$$L(G, 0) = \{a^m b^n \mid m > n \geqslant 1\}\ , \quad (8)$$

$$L(G, =, 0) = \{a^n b^n \mid n \geqslant 1\}\ . \quad (9)$$

Note that $L(G, 0)$ and $L(G, =, 0)$ are all CF languages in spite of the grammar G with type 3 rules only.

Example 3. In the above example let M be the cartesian product of Z (= set of integers), i.e. $M = Z \times Z$, and for each $(a,b), (c,d) \in Z \times Z$ let us define the operations $\times$ and + in M as

$$(a,b) \times (c,d) = (a \times c, b \times d)\ ,$$

$$(a,b) + (c,d) = (a + c, b + d)\ .$$

Then $M = Z \times Z$ forms a ring. Moreover let us define the order relation $>$ over M (= $Z \times Z$) as follows.

$$(a,b) > (c,d) \quad \text{iff} \quad a > c \quad \text{and} \quad b > d\ .$$

Productions are

(1) $S \to aS$ $\quad (+1, -1)$,

(2) $S \to aA$ $\quad (+1, -1)$,

(3) $A \to bA$ $\quad (-1, +1)$,

(4) $A \to b$ $\quad (-1, +1)$.

Then a string a^3b^2 is defined as

$$S \xrightarrow[1]{(+1,-1)} aS \xrightarrow[1]{(+1,-1)} a^2S \xrightarrow[2]{(+1,-1)}$$

$$a^3A \xrightarrow[3]{(-1,+1)} a^3bA \xrightarrow[4]{(-1,+1)} a^3b^2\ .$$

As there is no other derivation of a^3b^2, the weight of a^3b^2 is $\mu_G(a^3b^2) = (1, -1)$. Continuing in this manner we see that $\{(x, \mu_G(x))\}$ is

$$\{(a^i b^j, (i-j, j-i)) \mid i,j \geqslant 1\}\ .$$

From (5) we have the language with $\lambda = (-1, -1)$ as follows.

$$L(G, (-1,-1)) = \{a^n b^n \mid n \geqslant 1\}\ . \quad (10)$$

Note: From the definition of language of eq. (5) which is the most orthodox definition of language with threshold (or cut point) (cf. [9]), we cannot obtain the language $\{a^n b^n \mid n \geqslant 1\}$ from Example 3 [see (8)] but we can get the language $\{a^n b^n \mid n \geqslant 1\}$ from Example 4 [see (10)].

Next we shall construct the grammar with weights consisting of type 3 rules which can generate CS languages in the same way as Example 2 and 3.

Example 4. Let G be the type 3 formal grammar with weights as $G = (V_N, V_T, P, S, J, M, \mu)$. The weighting space M and the operations $\times$, $+$ are defined in the same manner as in Example 3. Let P be

(1) $S \to aS$ $(1,0)$,

(2) $S \to aA$ $(1,0)$,

(3) $A \to bA$ $(-1,1)$,

(4) $A \to bB$ $(-1,1)$,

(5) $B \to cB$ $(0,-1)$,

(6) $B \to c$ $(0,-1)$.

Then the derivation chain of a string, say, a^2bc is

$$S \underset{1}{\overset{(1,0)}{\longrightarrow}} aS \underset{2}{\overset{(1,0)}{\longrightarrow}} a^2A \underset{4}{\overset{(-1,1)}{\longrightarrow}} a^2bB \underset{6}{\overset{(0,-1)}{\longrightarrow}} a^2bc .$$

It is shown that there is no other derivation chain for a^2bc. Therefore we see that

$$\mu_G(a^2bc) = (1,0) + (1,0) + (-1,1) + (0,-1) = (1,0) .$$

In general we have

$$\{(a^ib^jc^k, (i-j, j-k)) \mid i, j, k \geq 1\} .$$

Hence let λ be $(0,0) \in M$, then from (5) and (6)

$$L(G, (0,0)) = \{a^m b^n c^p \mid m > n > p \geq 1\} , \tag{11}$$

$$L(G, =, (0,0)) = \{a^n b^n c^n \mid n \geq 1\} . \tag{12}$$

Note that the above languages are all CS languages.

Example 5. In Example 4 let M be the set of 2×2 matrices whose elements are integers ($\in Z$). For each $a, b, \cdots, h \in Z$ let us define the operations $\times$ and $+$ on M as

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a \times e & b \times f \\ c \times g & d \times h \end{bmatrix},$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}.$$

Then M forms a ring. The order relation $>$ over M is defined as

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} > \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

$$\text{iff} \quad a > e, b > f, c > g, \text{and } d > h .$$

Productions are given as follows.

(1) $S \to aS$ $\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$,

(2) $S \to aA$ $\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$,

(3) $A \to bA$ $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$,

(4) $A \to bB$ $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$,

(5) $B \to cB$ $\begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}$,

(6) $B \to c$ $\begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}$.

For example, a string a^2bc is obtained by the following derivation chain which is the only derivation chain for a^2bc.

$$S \underset{1}{\overset{\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}}{\longrightarrow}} aS \underset{2}{\overset{\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}}{\longrightarrow}} a^2A \underset{4}{\overset{\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}}{\longrightarrow}} a^2bB \underset{6}{\overset{\begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}}{\longrightarrow}} a^2bc.$$

Therefore we have

$$\mu_G(a^2bc) = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}.$$

In general $\{(x, \mu_G(x))\}$ is obtained as follows.

$$\left\{\left(a^ib^jc^k, \begin{bmatrix} i-j & j-k \\ -(i-j) & -(j-k) \end{bmatrix}\right) \middle| i, j, k \geq 1\right\}.$$

Therefore from (5) the language $L(G, \lambda)$ with $\lambda = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \in M$ is given as

$$L\left(G, \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}\right) = \{a^n b^n c^n | n \geqslant 1\} .$$

As mentioned above, we have seen that from several simple examples the generative powers of formal grammars with weights can be enhanced by incorporating the algebra systems with formal grammar systems.

More detailed studies on the properties of formal grammars with weights in which the weighting space is restricted to a particular algebra such as distributive lattices, more closely, the representative boolean lattice as given in fig. 1 will be presented in subsequent papers.

References

[1] M. Mizumoto, J. Toyoda and K. Tanaka, General formulation of formal grammars, Inform. Sc. 4 (1972) 87–100.
[2] A. Salomaa, Probabilistic and weighted grammars, Inform. and Control 15 (1969) 529–544.
[3] T. Huang and K.S. Fu, On stochastic context-free languages, Inform. Sc. 3 (1971) 201–224.
[4] E.T. Lee and L.A. Zadeh, Note on fuzzy languages, Inform. Sc. 1 (1969) 421–434.
[5] M. Mizumoto, J. Toyoda and K. Tanaka, N-fold fuzzy grammars, Inform. Sc. 5 (1973) 25–43.
[6] M. Mizumoto, Fuzzy Automata and Fuzzy grammars, Ph.D. diss., Osaka University, February 1971.
[7] M. Mizumoto, J. Toyoda and K. Tanaka, On fuzzy languages, Trans. Inst. Elect. Commun. Engrs. Japan 53-c (1970) 333–339 (in Japanese).
[8] J.E. Hopcroft and J.D. Ullman, Formal Languages and Their Relation to Automata (Addison-Wesley, Reading, Mass., 1969).
[9] A. Paz, Some aspects of probabilistic automata, Inform. and Control 9 (1966) 26–60.