A Note on Language Classes with Finite Elasticity

Makoto Kanazawa
Department of Cognitive and Information Sciences
Faculty of Letters, Chiba University
1–33 Yayoi-cho, Inage-ku
Chiba-shi, 263, Japan
kanazawa@cogsci.l.chiba-u.ac.jp

Abstract

It is proved that finite elasticity of language classes is preserved under the inverse image of a finite-valued relation, extending results of Wright's and of Moriyama and Sato's.

Keywords and Phrases: finite elasticity, identification in the limit, inductive inference, learnability.
Note: In preparing the present note, the author was sponsored by project NF 102/62–356 ('Structural and Semantic Parallels in Natural Languages and Programming Languages'), funded by the Netherlands Organization for the Advancement of Research (N.W.O.). The material presented here is based on Chapter 3 of my dissertation.

Let $\mathbf{S}$ be some set of objects. A subset $\mathcal{L}$ of pow($\mathbf{S}$) is said to have *infinite elasticity* if there exist an infinite sequence $\langle s_n \rangle_{n \in \mathsf{N}}$ of elements of $\mathbf{S}$ and an infinite sequence $\langle L_n \rangle_{n \in \mathsf{N}}$ of sets in $\mathcal{L}$ such that for all $n \in \mathsf{N}$,

$$s_n \notin L_n,$$

and

$$\{s_0, \ldots, s_n\} \subseteq L_{n+1}.$$

If $\mathcal{L}$ does not have infinite elasticity, it is said to have *finite elasticity*. The notion of finite elasticity was introduced by Wright (1989) in connection with inductive inference of formal languages from positive data (Gold 1967, Angluin 1980), where elements of $\mathbf{S}$ are strings over some finite alphabet, and subsets of $\mathrm{pow}(\mathbf{S})$ are classes of languages. The interest of finite elasticity lies in the fact that if $\{L_i\}_{i\in\mathsf{N}}$ is a *uniformly recursive* class of languages, then finite elasticity of $\{L_i\}_{i\in\mathsf{N}}$ implies that it is *identifiable in the limit from positive data*. (Gold 1967).[1] The main result of Wright 1989 is that finite elasticity is preserved under pointwise union: if $\mathcal{L}$ and $\mathcal{M}$ are two classes with finite elasticity, then the class $\{\, L \cup M \mid L \in \mathcal{L} \wedge M \in \mathcal{M} \,\}$ also has finite elasticity. Wright's proof of this result uses Ramsey's Theorem.[2] Recently, Moriyama and Sato 1993 shows that finite elasticity is preserved under many other operations as well, including pointwise concatenation and pointwise Kleene closure.

In this note, we prove a theorem on finite elasticity, which generalizes the essence of Wright's Theorem and many of Moriyama and Sato's results. The method of my proof is essentially the same as the one used by Moriyama and Sato, but these authors do not state their result in the general form given here.[3] The theorem has a number of applications; it is used extensively in my dissertation (Kanazawa 1994).

Let $\mathbf{S}$ and $\mathbf{U}$ be two (not necessarily distinct) sets of objects. A relation $R \subseteq \mathbf{S} \times \mathbf{U}$ is said to be *finite-valued* iff for every $s \in \mathbf{S}$, there are at most finitely many $u \in \mathbf{U}$ such that Rsu. If M is a subset of $\mathbf{U}$, define a subset $R^{-1}[M]$ of $\mathbf{S}$ by $R^{-1}[M] = \{\, s \mid \exists u(Rsu \wedge u \in M) \,\}$.

Theorem 1 *Let $\mathcal{M}$ be a subset of* $\mathrm{pow}(\mathbf{U})$ *that has finite elasticity, and let $R \subseteq \mathbf{S} \times \mathbf{U}$ be a finite-valued relation. Then $\mathcal{L} = \{\, R^{-1}[M] \mid M \in \mathcal{M} \,\}$ also has finite elasticity.*

Proof. Suppose that $\mathcal{L} = \{\, R^{-1}[M] \mid M \in \mathcal{M} \,\}$ has *in*finite elasticity. Then there is an infinite sequence of elements $s_0, s_1, s_2, \ldots$ of $\mathbf{S}$ and an infinite

[1] Moreover, as Kapur (1991) emphasizes, given the indexing of a uniformly recursive class of languages that has finite elasticity, one can synthesize a learning algorithm for that class.

[2] Unfortunately, Wright's original definition of finite elasticity was in error, and was later corrected by Motoki, Shinohara and Wright (1991).

[3] I came up with my proof in November 1993, before I became aware of Moriyama and Sato 1993. In their proof, Moriyama and Sato essentially reproduce the proof of König's Lemma, which I explicitly rely on in my proof.

sequence of sets $L_0, L_1, L_2, \ldots$ from $\mathcal{L}$ such that for each n, $s_n \notin L_n$ and $\{s_0, \ldots, s_n\} \subseteq L_{n+1}$. For each $n \in \mathsf{N}$, take an $M_n \in \mathcal{M}$ such that $L_n = R^{-1}[M_n]$. For each $k \in \mathsf{N}$, let

$$U_k = \{ \langle u_0, \ldots, u_k \rangle \mid Rs_0u_0 \wedge \cdots \wedge Rs_ku_k \wedge \exists n(\{u_0, \ldots, u_k\} \subseteq M_n) \}.$$

Note that each U_k is non-empty, and U_i and U_j are disjoint if $i \neq j$. Let

$$U = \bigcup_{k \in \mathsf{N}} U_k.$$

By the preceding remarks, U is infinite. U has the form of a tree: the mother of $\langle u_0, \ldots, u_k, u_{k+1} \rangle \in U$ is $\langle u_0, \ldots, u_k \rangle$, which is also in U. Since R is finite-valued, U is finitely branching. Since U is an infinite tree, by König's Lemma, U has an infinite branch. Let $u_0, u_1, u_2, \ldots$ be an infinite sequence of strings over Υ that corresponds to an infinite branch of U; i.e., $\langle u_0 \rangle, \langle u_0, u_1 \rangle, \langle u_0, u_1, u_2 \rangle, \ldots$ are the nodes on this branch. Note that $s_n \notin L_n$ implies

$$u_n \notin M_n. \tag{1}$$

For each n, let $f(n)$ be such that $\{u_0, \ldots, u_n\} \subseteq M_{f(n)}$ and for all $j < f(n)$, $\{u_0, \ldots, u_n\} \not\subseteq M_j$. By (1), $n < f(n)$ for all n. For each n, let $g(n) = f^n(0) = \underbrace{f(\ldots(f}_{n \text{ times}}(0))\ldots)$. Note that g is monotone increasing. We claim that

$$u_{g(0)}, u_{g(1)}, \ldots, u_{g(n)}, \ldots$$

and

$$M_{g(0)}, M_{g(1)}, \ldots, M_{g(n)}, \ldots$$

witness the infinite elasticity of $\mathcal{M}$. We have (1), so it is enough to observe that by the definition of g,

$$\{u_{g(0)}, \ldots, u_{g(n)}\} \subseteq M_{g(n+1)}$$

for all $n \in \mathsf{N}$. ∎

Note that an analogue of Theorem 1 does not hold of *ineffective identifiability in the limit from positive data* (Gold 1967), which is a property of language classes.[4] For instance, let $\mathsf{S} = \mathsf{N}$, the set of natural numbers, and take $\mathcal{M} = \{\mathsf{E}\} \cup \{\, \{1, 3, \ldots, 2n+1\} \mid n \in \mathsf{N} \,\}$, where E is the set of even numbers. $\mathcal{M}$ is (ineffectively) identifiable in the limit from positive data. The relation R defined by $Rxy \Leftrightarrow y = 2x \vee y = 2x+1$ is finite-valued. But $\mathcal{L} = \{\, R^{-1}[M] \mid M \in \mathcal{M} \,\} = \{\mathsf{N}\} \cup \{\, \{0, 1, \ldots, n\} \mid n \in \mathsf{N} \,\}$ is not ineffectively identifiable in the limit from positive data.

Below we list some applications of Theorem 1.

Example 2 Let $\mathcal{M}$ and $\mathcal{N}$ be two classes of languages with finite elasticity. Let $\mathtt{0}$ and $\mathtt{1}$ be symbols that do not appear in $\mathcal{M}$ and $\mathcal{N}$, and let $M \uplus N = \{\, u\mathtt{0} \mid u \in M \,\} \cup \{\, v\mathtt{1} \mid v \in N \,\}$.

It is easy to see that the class $\{\, M \uplus N \mid M \in \mathcal{M} \wedge N \in \mathcal{N} \,\}$ also has finite elasticity. For, suppose that it has *in*finite elasticity. Then there exist an infinite sequence of strings $w_0, w_1, w_2, \ldots$ and infinite sequences of languages $M_0, M_1, M_2, \ldots$ and $N_0, N_1, N_2, \ldots$ from $\mathcal{M}$ and $\mathcal{N}$ respectively, such that for each $n \in \mathsf{N}$,

$$w_n \notin M_n \uplus N_n,$$

and

$$\{w_0, \ldots, w_n\} \subseteq M_{n+1} \uplus N_{n+1}.$$

There must be an infinite subsequence $w_{i_0}, w_{i_1}, w_{i_2}, \ldots$ of $w_0, w_1, w_2, \ldots$ such that either (i) for all $n \in \mathsf{N}$, w_{i_n} is of the form $u\mathtt{0}$, or (ii) for all $n \in \mathsf{N}$, w_{i_n} is of the form $v\mathtt{1}$. Assume (i), and let $w_{i_n} = u_{i_n}\mathtt{0}$. For each $n \in \mathsf{N}$, since $u_{i_n}\mathtt{0} \notin M_{i_n} \uplus N_{i_n}$, $u_{i_n} \notin M_{i_n}$. Since $\{u_{i_0}\mathtt{0}, \ldots, u_{i_n}\mathtt{0}\} \subseteq M_{i_{n+1}} \uplus N_{i_{n+1}}$, $\{u_{i_0}, \ldots, u_{i_n}\} \subseteq M_{i_{n+1}}$. Thus,

$$u_{i_0}, u_{i_1}, u_{i_2}, \ldots, u_{i_n}, \ldots$$

and

$$M_{i_0}, M_{i_1}, M_{i_2}, \ldots, M_{i_n}, \ldots$$

witness the infinite elasticity of $\mathcal{M}$, contradicting the assumption. The case where (ii) holds is similar.

Let R be the finite-valued relation such that Rsw iff $w = s\mathtt{0}$ or $w = s\mathtt{1}$. Let $\mathcal{L} = \{\, R^{-1}[M \uplus N] \mid M \in \mathcal{M} \wedge N \in \mathcal{N} \,\}$. Note that $\mathcal{L} = \{\, M \cup N \mid M \in$

[4]Ineffectively identifiability is just like identifiability, except that the requirement that the learning function must be effectively computable is dropped. For ineffective identifiability, the indexing of the give language class $\{L_i\}_{i \in \mathsf{N}}$ is irrelevant; thus ineffectively identifiability is purely an extensional, set-theoretic property, just like finite elasticity.

$\mathcal{M} \wedge N \in \mathcal{N}\,\}$. Since $\{\, M \uplus N \mid M \in \mathcal{M} \wedge N \in \mathcal{N}\,\}$ has finite elasticity, by Theorem 1, so does $\mathcal{L}$. Thus, Wright's (1989) theorem follows as a special case of Theorem 1. ▪▪▪

Example 3 Let $\mathcal{M}$ be a class of languages over Υ, and let $h\colon \Upsilon^* \to \Sigma^*$ be a non-erasing homomorphism; i.e., h is the unique homomorphism that extends some $h_0\colon \Upsilon \to \Sigma^+$. The relation $R \subseteq \Sigma^* \times \Upsilon^*$ defined by $Rsu \Leftrightarrow s = h(u)$ is finite-valued, for, if $s = h(u)$, $|u| \leq |s|$, where $|v|$ is the length of v. Let $\mathcal{L} = \{\, h[M] \mid M \in \mathcal{M}\,\} = \{\, R^{-1}[M] \mid M \in \mathcal{M}\,\}$. By Theorem 1, $\mathcal{L}$ has finite elasticity if $\mathcal{M}$ does. ▪▪▪

Example 4 Let $\mathcal{M}$ be a class of languages with finite elasticity. Then, the class $\mathcal{L}$ of permutation closures of languages in $\mathcal{M}$ also has finite elasticity. In general, take any relation $R \subseteq \Sigma^* \times \Sigma^*$ such that Rsu only if $|s| = |u|$. If a class $\mathcal{M}$ of languages over Σ has finite elasticity, then $\mathcal{L} = \{\, R^{-1}[M] \mid M \in \mathcal{M}\,\}$ has finite elasticity. ▪▪▪

Example 5 Assume that an ordered pair $\langle u, v\rangle$ of strings u and v is encoded as a string $\texttt{<}u\texttt{,}v\texttt{>}$, where $\texttt{<}$, $\texttt{,}$, $\texttt{>}$ are new symbols. If M and N are languages, let $M \times N = \{\, \langle u, v\rangle \mid u \in M \wedge v \in N\,\}$.

Let $\mathcal{M}$ and $\mathcal{N}$ be classes of languages with finite elasticity. It is easy to see that the class $\mathcal{M} \times \mathcal{N} = \{\, M \times N \mid M \in \mathcal{M} \wedge N \in \mathcal{N}\,\}$ also has finite elasticity. For, suppose not. Then there exist infinite sequences of strings $u_0, u_1, u_2, \ldots$ and $v_0, v_1, v_2, \ldots$ and infinite sequences of languages $M_0, M_1, M_2, \ldots$ and $N_0, N_1, N_2, \ldots$ such that for each $n \in \mathsf{N}$,

$$\langle u_n, v_n\rangle \notin M_n \times N_n \tag{2}$$

and

$$\{\langle u_0, v_0\rangle, \ldots, \langle u_n, v_n\rangle\} \subseteq M_{n+1} \times N_{n+1}. \tag{3}$$

By (2), for each $n \in \mathsf{N}$, either $u_n \notin M_n$ or $v_n \notin N_n$. Thus, either for infinitely many n, $u_n \notin M_n$, or for infinitely many n, $v_n \notin N_n$. Assume the former, the latter being symmetric. Thus, there is an infinite subsequence $u_{i_0}, u_{i_1}, u_{i_2}, \ldots$ of $u_0, u_1, u_2, \ldots$ such that for each n, $u_{i_n} \notin M_{i_n}$. By (3), $\{u_{i_0}, \ldots, u_{i_n}\} \subseteq M_{i_{n+1}}$. This means that $u_{i_0}, u_{i_1}, u_{i_2}, \ldots$ and $M_{i_0}, M_{i_1}, M_{i_2}, \ldots$ witness the infinite elasticity of $\mathcal{M}$, contradicting the assumption.

Let $\mathrm{shuffle}(M, N) = \{\, w \mid \exists u_0, u_1, \ldots, u_n, v_0, v_1, \ldots, v_n (w = u_0 v_0 u_1 v_1 \ldots u_n v_n \wedge u_0 u_1 \ldots u_n \in M \wedge v_0 v_1 \ldots v_n \in N)\,\}$. $\mathrm{shuffle}(M, N)$ is the set of strings that can be obtained by interleaving a string from M and a string from N. One can now apply Theorem 1 to show that the class $\mathcal{L} = \{\, \text{shuffle}(M, N) \mid M \in \mathcal{M} \wedge N \in \mathcal{N} \,\}$ has finite elasticity. To see this, take the relation R such that Rsw if and only if for some $u_0, u_1, \ldots, u_n, v_0, v_1, \ldots, v_n$, $s = u_0 v_0 u_1 v_1 \ldots u_n v_n$ and $w = \langle u_0 u_1 \ldots, u_n, v_0 v_1 \ldots v_n \rangle$. R is a finite-valued relation, for, if $R\, s\, \langle u, v \rangle$, then $|s| = |uv|$. It remains to note that $\mathcal{L} = \{\, R^{-1}[M \times N] \mid M \in \mathcal{M} \wedge N \in \mathcal{N} \,\}$. ■ ■ ■

Example 6 If s is a string over Σ, let $\frac{1}{2}s$ be that initial segment of s such that $|\frac{1}{2}s| = \lceil \frac{1}{2}|s| \rceil$. Let $\frac{1}{2}L = \{\, \frac{1}{2}s \mid s \in L \,\}$. If $\mathcal{M}$ is a class of languages with finite elasticity, then $\mathcal{L} = \{\, \frac{1}{2}M \mid M \in \mathcal{M} \,\}$ also has finite elasticity. To see this, take the finite-valued relation R such that Rsu if and only if $s = \frac{1}{2}u$. ■ ■ ■

Example 7 Let $\mathcal{M}$ be a class of ϵ-free languages with finite elasticity. Let $\mathcal{M}^n = \{\, M_1 \times \cdots \times M_n \mid M_1, \ldots, M_n \in \mathcal{M} \,\}$, where $M_1 \times \cdots \times M_n = \{\, \langle u_1, \ldots, u_n \rangle \mid u_1 \in M_1 \wedge \cdots \wedge u_n \in M_n \,\}$, assuming a suitable encoding of ordered n-tuples. As in the case of Example 5, it is easy to see that $\mathcal{M}^n$ has finite elasticity. Since every language in $\mathcal{M}^i$ is disjoint from every language in $\mathcal{M}^j$ if $i \neq j$, $\bigcup_n \mathcal{M}^n$ must also have finite elasticity. Let R be such that Rsu if and only if for some non-empty strings $s_1, \ldots, s_n$ such that $s = s_1 \ldots s_n$, $u = \langle s_1, \ldots, s_n \rangle$. Then R is a finite-valued relation. Let $\mathcal{L} = \{\, R^{-1}[M] \mid M \in \bigcup_n \mathcal{M}^n \,\}$. $\mathcal{L}$ consists of languages that are concatenations of languages from $\mathcal{M}$. By Theorem 1, $\mathcal{L}$ has finite elasticity. (This does not in general hold without the assumption that $\mathcal{M}$ consists of ϵ-free languages. Let $M = \{\epsilon, \mathsf{a}\}$. Then $\{M\}^n = \{M^n\}$, where $M^n = \underbrace{M \times \cdots \times M}_{n \text{ times}}$. $R^{-1}[M^n] = \{\epsilon, \mathsf{a}, \mathsf{aa}, \ldots, \mathsf{a}^n\}$. $\{\, R^{-1}[M^n] \mid n \in \mathsf{N} \,\} = \{\, R^{-1}[N] \mid N \in \bigcup_n \{M\}^n \,\}$ has *in*finite elasticity.) ■ ■ ■

Example 8 Let $\mathcal{G}$ be a class of context-free grammars over Σ that do not contain unit productions (production rules of the form $A \to B$) or ϵ productions ($A \to \epsilon$). Let $\mathcal{M}$ be the class consisting of the sets of *skeletal phrase structures* (Levy and Joshi 1978) generated by the grammars in $\mathcal{G}$. Let $\mathcal{L}$ be the class of languages generated by grammars in $\mathcal{G}$ in the usual sense. Then, if $\mathcal{M}$ has finite elasticity, so does $\mathcal{L}$. To see this, take the finite-valued relation $s = \text{yield}(T)$ between strings and skeletal phrase structures of the appropriate sort. ■ ■ ■

See Moriyama and Sato (1993) and Kanazawa (1994) for further examples of easy consequences of Theorem 1.

References

Angluin, Dana. 1980b. Inductive inference of formal languages from positive data. *Information and Control* **45**, 117–135.

Gold, E. M. 1967. Language identification in the limit. *Information and Control* **10**, 447–474.

Kanazawa, Makoto. 1994. *Learnable Classes of Categorial Grammars*. Ph.D. dissertation, Department of Linguistics, Stanford University. Available as ILLC Dissertation Series 1994–8, ILLC, University of Amsterdam (`illc@fwi.uva.nl`).

Kapur, Shyam. 1991. *Computational Learning of Languages*. Ph.D. dissertation, Cornell University. Available as Technical Report 91–1234, Department of Computer Science, Cornell University.

Levy, Leon S. and Aravind K. Joshi. 1978. Skeletal structural descriptions. *Information and Control* **39**, 192–211.

Moriyama, Takashi and Masako Sato. 1993. Properties of language classes with finite elasticity. In K. P. Jankte, S. Kobayashi, E. Tomita, and T. Yokomori, eds., *Algorithmic Learning Theory*. Proceedings, 1993. Lecture Notes in Artificial Intelligence 744. Berlin: Springer.

Motoki, T., T. Shinohara, and K. Wright. 1991. The correct definition of finite elasticity: Corrigendum to identification of unions. In *The Fourth Annual Workshop on Computational Learning Theory*, p. 375. San Mateo, Calif.: Morgan Kaufmann.

Wright, K. 1989. Identification of unions of languages drawn from an identifiable class. In *The 1989 Workshop on Computational Learning Theory*, pp. 328–333. San Mateo, Calif.: Morgan Kaufmann.