Supplementary Material

I. Definitions

Let Σ be a finite set of constant symbols, e.g. $\{a, b, c, \dots\}$. Let Γ be a finite set of variable symbols, e.g. $\{X, Y, Z, \dots\}$. Let $P^n (n \geq 0)$ be a finite set of n-ary predicate symbols, and $P = \bigcup_{i \geq 0} P^i$. We use $p(t_1, \dots, t_k)$ to denote a k-ary predicate, where $p \in P^k, t_i \in \Sigma \cup \Gamma$. Let P, Q be two predicates. $\phi(P)$ is the arity of P. P and Q are identical if they share the same predicate symbol and argument list, written as $P \equiv Q$. A predicate p is called grounded if all arguments (i.e. $t_1, \dots$) are constants. The above definitions do not break the those defined in first-order predicate logic.

Definition 1 (Relational Database, RDB)**.** *A relational database is a finite set of grounded predicates.*

Formally, the pattern induced in SInc is a first-order Horn rule:

$$Q \leftarrow P_1 \land P_2 \land \cdots \land P_n$$

where $Q, P_1, \dots, P_n$ are predicates (arguments of which are from $\Sigma \cup \Gamma$) and there is no negation of the predicates in the rule. $\psi(r)$ denotes the number of different variables in r. According to some rule r, Q is entailed by $P_1 \land \cdots \land P_n$ if P_i are all true, that is $(\bigwedge_i P_i) \land (Q \leftarrow P_i) \models Q$. Thus by binding the variables in the entailment, the grounded predicate Q' is entailed by grounded predicates P'_i w.r.t. the rule r if P'_i is in the RDB $\mathcal{D}$, written as $\{P'_i\} \models_r Q'$. Let $\mathcal{S}, \mathcal{T}$ be sets of grounded predicates, $\mathcal{P}$ be a set of inference rules, $\mathcal{S} \models_{\mathcal{P}} \mathcal{T}$ if $\forall T \in \mathcal{T}, \exists \mathcal{S}' \subseteq \mathcal{S}, r \in \mathcal{P}$, such that $\mathcal{S}' \models_r T$. Suppose $\mathcal{S} \models_r T$. If $T \in \mathcal{D}$, T is said to be positively entailed by $\mathcal{S}$ w.r.t. r; otherwise, T is negatively entailed.

Definition 2 (Limited Variable (LV), Unlimited Variable (UV))**.** *A variable is unlimited in some inference rule r if there is only one argument in r that is assigned to it. Each unique UV is referred by a '?'. A variable is limited in r if there exist at least two arguments in r that are assigned to it.*

For example, the following two rules are identical: X is a LV; Y and Z are UVs.

$$q(X, Y) \leftarrow p(X, Z)$$
$$q(X, ?) \leftarrow p(X, ?)$$

Definition 3 (Length of Horn rules)**.** *The length of rule r is:*

$$|r| = \left(\sum_{P \in r} \phi(P) \right) - \psi(r)$$

Definition 4 (Semantic Inductive Compression, SIC)**.** *Let $\mathcal{D}$ be a relational database. The compression on $\mathcal{D}$ is a triple $(\mathcal{P}, \mathcal{R}, \mathcal{I}^-)$ with minimal size, where $\mathcal{P}$ is a set of inference rules, both $\mathcal{R}$ and $\mathcal{I}^-$ are sets of grounded predicates. $\mathcal{D}, \mathcal{P}, \mathcal{R}, \mathcal{I}^-$ satisfies:*

- $\mathcal{R} \subseteq \mathcal{D}$
- $\mathcal{R} \models_{\mathcal{P}} (\mathcal{D} \setminus \mathcal{R}) \cup \mathcal{I}^-$
- $\forall e \notin \mathcal{D} \cup \mathcal{I}^-, \forall r \in \mathcal{P}, \mathcal{R} \not\models_r e$

The size of $(\mathcal{P}, \mathcal{R}, \mathcal{I}^-)$ is $|\mathcal{P}| + |\mathcal{R}| + |\mathcal{I}^-|$. $|\mathcal{R}|$ is the number of predicates in $\mathcal{R}$, and so be $|\mathcal{I}^-|$. $|\mathcal{P}|$ is defined as the sum of lengths of all rules in it.

Definition 5 (SIC, decision version)**.** *Let $\mathcal{D}$ be a relational database and k be a positive integer. The compression on $\mathcal{D}$ given k is to determine whether there is a triple $(\mathcal{P}, \mathcal{R}, \mathcal{I}^-)$ with size no larger than k that satisfies:*

- $\mathcal{R} \subseteq \mathcal{D}$
- $\mathcal{R} \models_{\mathcal{P}} (\mathcal{D} \setminus \mathcal{R}) \cup \mathcal{I}^-$
- $\forall e \notin \mathcal{D} \cup \mathcal{I}^-, \forall r \in \mathcal{P}, \mathcal{R} \not\models_r e$

Let r be some inference rule. If for some $\mathcal{R} \subseteq \mathcal{D}$, $\mathcal{R} \models_r \mathcal{D} \setminus \mathcal{R}$, r separates $\mathcal{D}$ into two parts: $\mathcal{R}$ and $\mathcal{D} \setminus \mathcal{R}$, where the latter can be removed from $\mathcal{D}$. Let $\mathcal{I}^-$ be the set of all negatively inferred predicates from $\mathcal{R}$ w.r.t. r, $(\{r\}, \mathcal{R}, \mathcal{I}^-)$ is a candidate solution (not necessarily minimum) to SIC. The size reduced by r is:

$$\Delta(r) = |\mathcal{D}| - (|r| + |\mathcal{R}| + |\mathcal{I}^-|) = |\mathcal{D} \setminus \mathcal{R}| - |r| - |\mathcal{I}^-|$$

II. Problem Complexity

Theorem 6. *SIC is in NP for fixed $\mathcal{P}$.*

Proof. According to the definition, $\mathcal{P}$ is a Datalog program. Let $\mathcal{A}$ be the evaluation result of $\mathcal{P}$ on $\mathcal{R}$. $(\mathcal{P}, \mathcal{R}, \mathcal{I}^-)$ is a valid solution if the following are all true:

- $|\mathcal{P}| + |\mathcal{R}| + |\mathcal{I}^-| \leq k$
- $\mathcal{R} \subseteq \mathcal{D}$
- $\mathcal{A} \cap \mathcal{D} \supseteq \mathcal{D} \setminus \mathcal{R}$
- $\mathcal{A} \setminus \mathcal{D} = \mathcal{I}^-$

All above comparison can be finished in polynomial time w.r.t. $\mathcal{D}, \mathcal{R}, \mathcal{I}^-, \mathcal{P}$ and $\mathcal{A}$. Moreover, $\mathcal{A}$ is computable in polynomial time of $\mathcal{R}$ for fixed $\mathcal{P}$ [1]. Therefore, for fixed $\mathcal{P}$, the overall verification of the solution is in polynomial time w.r.t. $\mathcal{D}$, $\mathcal{R}$ and $\mathcal{I}^-$. $\square$

Definition 7 (Vertex Cover Problem)**.** *Let $\mathcal{G}_{vc} = \langle \mathcal{V}_{vc}, \mathcal{E}_{vc} \rangle$ be an undirected graph. Let k be a positive integer. A vertex cover $\mathcal{V}_c$ of $\mathcal{G}_{vc}$ is a subset of $\mathcal{V}_{vc}$ such that $(u, v) \in \mathcal{E}_{vc} \implies u \in \mathcal{V}_c \lor v \in \mathcal{V}_c$. The vertex cover problem is to determine whether there is a vertex cover $\mathcal{V}_c$ of $\mathcal{G}_{vc}$ s.t. $|\mathcal{V}_c| \leq k$.*

Hardness of the semantic compression can be proved by reducing the vertex cover problem to SIC. Let $\mathcal{G}_{vc} = \langle \mathcal{V}_{vc}, \mathcal{E}_{vc} \rangle$ be the graph in the vertex cover problem and $n = |\mathcal{V}_{vc}|$, $m = |\mathcal{E}_{vc}|$. Let m_v be the number of edges connected to vertex v. By the following settings we create a relational database aligning with $\mathcal{G}_{vc}$:

- Let $edge$ be a unary relation in $\mathcal{D}$ for edges and v be a unary relation for each $v \in \mathcal{V}_{vc}$;

- For each $(u, v) \in \mathcal{E}_{vc}$, add three predicates to $\mathcal{D}$: $edge(e)$, $u(e)$, $v(e)$;
- Add $m + 2$ redundant predicates: $f_1(c_1)$, $\ldots$, $f_{m+2}(c_{m+2})$.

The number of predicates, predicate symbols and constant symbols in $\mathcal{D}$ are $4m+2$, $m+n+3$ and $2m+2$. Therefore, the reduction can be done in linear time.

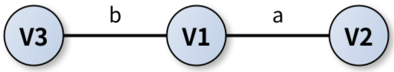

Fig. 1. Vertex Cover Example

For example, Figure 1 shows a graph with three vertices and two edges. The corresponding database contains the following predicates:

- $edge(a)$, $edge(b)$
- $v_1(a)$, $v_2(a)$, $v_1(b)$, $v_3(b)$
- $f_1(c_1), \ldots, f_4(c_4)$

Lemma 8. *If $\mathcal{G}_{vc}$ can be covered by $\mathcal{V}_c$ s.t. $|\mathcal{V}_c| \leq k$, there is a solution $(\mathcal{P}, \mathcal{R}, \mathcal{I}^-)$ for the corresponding compression problem s.t. $|\mathcal{P}| + |\mathcal{R}| + |\mathcal{I}^-| \leq 3m + 2 + k$*

Proof. Let $\mathcal{P} = \{edge(X) \leftarrow v(X) | v \in \mathcal{V}_c\}$, $\mathcal{E} = \{edge(e) | e \in \mathcal{E}_{vc}\}$, $\mathcal{V} = \{v(e) | \exists u, e(u, v) \in \mathcal{E}_{vc}$ or $e(v, u) \in \mathcal{E}_{vc}\}$. For each $edge(e) \in \mathcal{E}$, $\exists r \in \mathcal{P}$, s.t. $\mathcal{V} \models_r edge(e)$ and there is no counter example as there is some $v \in \mathcal{V}_c$ that covers the corresponding edge. Moreover, $\mathcal{V} \subseteq \mathcal{D}$. Therefore, $(\mathcal{P}, \mathcal{D} \setminus \mathcal{E}, \varnothing)$ is a valid solution to SIC. $|\mathcal{P}| + |\mathcal{D} \setminus \mathcal{E}| + |\varnothing| = |\mathcal{V}_c| \cdot 1 + 4m + 2 - m \leq 3m + 2 + k$. ☐

Lemma 9. *Considering all possible forms of a single inference rule, rules of the following form brings the most reduction on the database, which is $m_v - 1$:*

$$edge(X) \leftarrow v(X), m_v \geq 1 \tag{1}$$

Proof. Let r_v refer to the above rule. According to the construction of $\mathcal{D}$ and other definitions, the size reduced by a single rule r_v is $m_v - |r_v| = m_v - 1 \geq 0$ as there is no counter example introduced by r_v. The reduction of the other forms of rules are:

- $\Delta(edge(?) \leftarrow true) = m - (m + 2) = -2$;
- $\Delta(v(?) \leftarrow true) = m_v - (2m + 2 - m_v) = 2m_v - 2m - 2 \leq -2$;
- $\Delta(f_i(?) \leftarrow true) = 1 - (2m + 1) = -2m \leq 0$;
- $\Delta(v(X) \leftarrow edge(X)) = m_v - 1 - m \leq -1$;
- $\Delta(v(X) \leftarrow u(X)) = b - 1 - (m_u - b) \leq 0$, where $b = [(u, v) \in \mathcal{E}_{vc} \vee (v, u) \in \mathcal{E}_{vc}]$, further more, when $b = 1$, $m_u \geq 1$, otherwise, $m_u \geq 0$;
- Given that only argument $v.0$ and $edge.0$ share some of constant symbols, any other longer forms of rules present no larger Δ than r_i above.

☐

Lemma 10. *If $(\mathcal{P}, \mathcal{R}, \mathcal{I}^-)$ is a solution to SIC and $|\mathcal{P}| + |\mathcal{R}| + |\mathcal{I}^-| \leq k$, there is another solution $(\mathcal{P}', \mathcal{R}', \varnothing)$ to SIC s.t.:*

- *$|\mathcal{P}'| + |\mathcal{R}'| \leq k$*
- *Rules in $\mathcal{P}'$ are all under the form of Rule (1);*
- *Let $\mathcal{E} = \{edge(e) | e \in \mathcal{E}_{vc}\}$, $\mathcal{R}' = \mathcal{D} \setminus \mathcal{E}$;*

Proof. Lemma 9 indicates that if some rule in $\mathcal{P}$ is not under the form of Rule (1), they can be simply removed or replaced with some r_v under the form of Rule (1) s.t. the size of compression does not increase, yielding $\mathcal{P}''$. All rules in $\mathcal{P}''$ are under the form of Rule (1). If $\exists edge(e_1), \ldots, edge(e_l) \in \mathcal{E}$ that are not inferable w.r.t. $\mathcal{P}''$, at most l rules under the form of Rule (1) that infer all of these predicates can be added to the hypothesis set, yielding $\mathcal{P}'$. The size of the compression does not increase. After the above procedure, all edges are inferred by some vertex, thus $\mathcal{R}' = \mathcal{D} \setminus \mathcal{E}$. Moreover, by $\mathcal{P}'$, there is no counter example in the solution. ☐

Lemma 11. *If $(\mathcal{P}, \mathcal{R}, \mathcal{I}^-)$ is a solution to SIC and $|\mathcal{P}| + |\mathcal{R}| + |\mathcal{I}^-| \leq 3m + 2 + k$, there is a $\mathcal{V}_c$ that covers $\mathcal{E}_{vc}$ and $|\mathcal{V}_c| \leq k$*

Proof. By Lemma 10, there is a solution $(\mathcal{P}', \mathcal{R}', \varnothing)$ with size no larger than $3m + 2 + k$. Let $\mathcal{V}_c = \{v | edge(X) \leftarrow v(X) \in \mathcal{P}'\}$. Vertices in $\mathcal{V}_c$ cover all edges as the rules in $\mathcal{P}'$ infer all edges in $\mathcal{D}$. $|\mathcal{V}_c| = |\mathcal{P}'| \leq 3m + 2 + k - |\mathcal{R}'| = k$. ☐

Theorem 12. *SIC is NP-Hard.*

Proof. By Lemma 8 and 11, the vertex cover problem has solution w.r.t. a positive integer k iff the reduced SIC problem has solution w.r.t. $3m + 2 + k$. According to the reduction setting, the vertex cover problem can be polynomially reduced to SIC. Therefore, SIC is NP-Hard. ☐

Theorem 13. *SIC is NP-Complete for fixed $\mathcal{P}$.*

Proof. According to Theorem 6 and 12, SIC is in NP for fixed hypothesis and is also NP-Hard. Therefore, SIC is NP-Complete. ☐

III. Greedy Algorithm for FVS

Detailed algorithms for FVS is shown in Algorithm 1.

Algorithm 1 FVS

Input: Dependency Graph $\mathcal{G}$

Output: Set of Vertices that cover every cycle in $\mathcal{G}$

$V' \leftarrow \varnothing$
$SCCs \leftarrow$ strongly connected components in $\mathcal{G}$
for each $SCC \in SCCs$ **do**
  while there are edges in SCC **do**
    $v \leftarrow$ vertex in SCC that has maximum in-degree $\times$ out-degree
    $V' \leftarrow V' \cup \{v\}$
    remove v and its adjacent edges
    while there exists $v' \in SCC$ such that its in-degree or out-degree is zero **do**
      remvoe v' and its adjacent edges
    end while
  end while
end for
return V'

IV. Detailed Experimental Results

Table I to IV show detailed results of compression enhancement comparison among SInC, KGist, AMIE and Gzip. Dataset names are listed in short: E(Elti), D(Dunur), DBf(DBpedia.factbook), Fs(Family.simple), Fm(Family.medium), U(UMLS), FB(FB15K), WN(WN18), N(NELL). "-E" means without enhancement and "+E" stands for compression ratios with enhancement.

References

[1] E. Dantsin, T. Eiter, G. Gottlob, and A. Voronkov, "Complexity and expressive power of logic programming," *ACM Computing Surveys (CSUR)*, vol. 33, no. 3, pp. 374–425, 2001.

TABLE I
ENHANCEMENT EFFECT - SINC

	Dataset	E	D	S	DBf	Fs	Fm	U	FB	WN	N
	Reduc. Ratio (%)	33.79	66.80	69.61	51.94	35.72	35.90	26.77	61.40	44.63	59.92
LZ4	-E	41.04	31.32	19.42	36.60	47.37	43.61	18.66	33.42	38.86	28.98
	+E	18.02	26.97	17.20	19.76	19.01	16.55	6.71	34.76	19.88	31.18
	Enh.	43.90	86.10	88.56	53.98	40.13	37.94	35.98	104.00	51.17	107.59
Lzop	-E	36.13	27.05	17.83	34.01	39.16	35.97	17.84	34.92	40.13	29.80
	+E	16.99	23.58	17.59	20.03	18.14	13.00	6.81	35.71	20.18	31.48
	Enh.	47.01	87.17	98.69	58.88	46.32	36.16	38.16	102.28	50.29	105.65
Pixz	-E	22.78	15.12	5.53	21.15	15.34	11.22	7.80	16.94	15.39	12.28
	+E	11.91	13.81	9.02	12.73	11.19	8.09	3.24	14.50	7.76	10.46
	Enh.	52.30	91.33	163.04	60.18	72.97	72.09	41.57	85.60	50.43	85.16
Gzip	-E	22.97	16.87	11.28	23.50	19.51	17.07	9.35	22.38	24.00	17.86
	+E	11.16	14.18	10.19	13.56	11.15	9.00	3.59	21.52	12.05	17.35
	Enh.	48.58	84.02	90.37	57.68	57.17	52.70	38.43	96.17	50.20	97.17
Bzip2	-E	18.77	12.46	7.98	19.49	14.47	13.19	5.33	15.92	15.87	13.64
	+E	10.54	10.62	6.34	12.18	8.08	5.89	2.39	13.80	7.76	11.36
	Enh.	56.15	85.23	79.42	62.53	55.87	44.67	44.86	86.71	48.91	83.22
7zip	-E	23.69	15.71	6.23	21.37	17.00	11.61	8.76	17.02	16.16	13.10
	+E	13.09	14.40	9.16	13.26	13.04	8.49	3.40	14.69	8.15	11.02
	Enh.	55.25	91.71	147.01	62.02	76.71	73.12	38.84	86.30	50.41	84.11
	Best $R_{\text{SINC}\cdot i}$	**10.54**	**10.62**	**6.34**	**12.18**	**8.08**	**5.89**	**2.39**	**13.80**	**7.76**	**10.46**
	Best $E_{\text{SINC}/i}$	**43.90**	**84.02**	**79.42**	**53.98**	**40.13**	**36.16**	**35.98**	**85.60**	**48.91**	**83.22**
	Best $E^{-1}_{\text{SINC}/i}$	**31.19**	**15.90**	**9.11**	**23.46**	**22.63**	**16.41**	**8.94**	**22.48**	**17.39**	**17.46**
	Best Enhancement Efficiency	**76.97**	**79.51**	**87.65**	**96.22**	**89.00**	**99.29**	**74.40**	**71.74**	**91.25**	**72.00**

TABLE II
ENHANCEMENT EFFECT - KGIST

	Dataset	E	D	S	DBf	Fs	Fm	U	FB	WN	N
	Reduc. Ratio (%)	-	-	-	175.78	299.36	257.41	68.40	192.92	135.28	157.82
LZ4	-E	41.04	31.32	19.42	36.60	47.37	43.61	18.66	33.42	38.86	14.08
	+E	-	-	-	119.80	207.30	170.39	46.57	153.72	119.49	99.02
	Enh.	-	-	-	327.27	437.64	390.71	249.57	459.94	307.48	703.24
Lzop	-E	36.13	27.05	17.83	34.01	39.16	35.97	17.84	34.92	40.13	14.48
	+E	-	-	-	116.37	202.01	165.62	48.72	121.15	99.05	84.64
	Enh.	-	-	-	342.16	515.88	460.49	273.19	346.96	246.81	584.63
Pixz	-E	22.78	15.12	5.53	21.15	15.34	11.22	7.80	16.94	15.39	5.97
	+E	-	-	-	52.97	96.19	65.65	13.43	35.39	29.44	32.71
	Enh.	-	-	-	250.45	627.03	584.97	172.15	208.84	191.27	548.09
Gzip	-E	22.97	16.87	11.28	23.50	19.51	17.07	9.35	22.38	24.00	8.68
	+E	-	-	-	87.08	144.24	115.26	29.76	76.87	66.48	57.82
	Enh.	-	-	-	370.49	739.43	675.04	318.21	343.48	277.04	666.38
Bzip2	-E	18.77	12.46	7.98	19.49	14.47	13.19	5.33	15.92	15.87	6.63
	+E	-	-	-	71.64	126.97	95.27	22.85	58.14	50.67	52.93
	Enh.	-	-	-	367.65	877.51	722.30	428.35	365.20	319.23	798.28
7zip	-E	23.69	15.71	6.23	21.37	17.00	11.61	8.76	17.02	16.16	6.37
	+E	-	-	-	54.17	99.40	66.11	13.47	35.43	29.51	32.15
	Enh.	-	-	-	253.46	584.76	569.19	153.77	208.15	182.59	505.09
	Best $R_{\text{KGIST}\cdot i}$	**-**	**-**	**-**	**52.97**	**96.19**	**65.65**	**13.43**	**35.39**	**29.44**	**32.15**
	Best $E_{\text{KGIST}/i}$	**-**	**-**	**-**	**250.45**	**437.64**	**390.71**	**153.77**	**208.15**	**182.59**	**505.09**
	Best $E^{-1}_{\text{KGIST}/i}$	**-**	**-**	**-**	**30.14**	**32.13**	**25.50**	**19.63**	**18.34**	**21.76**	**20.37**
	Best Enhancement Efficiency	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**

TABLE III
ENHANCEMENT EFFECT - AMIE

	Dataset	E	D	S	DBf	Fs	Fm	U	FB	WN	N
	Reduc. Ratio (%)	-	-	-	81.60	87.79	83.36	65.78	53.69	58.77	57.98
LZ4	-E	41.04	31.32	19.42	36.60	47.37	43.61	18.66	33.42	38.86	14.08
	+E	-	-	-	30.91	43.70	39.73	15.80	30.71	27.16	31.49
	Enh.	-	-	-	84.43	92.25	91.09	84.69	91.88	69.88	108.67
Lzop	-E	36.13	27.05	17.83	34.01	39.16	35.97	17.84	34.92	40.13	14.48
	+E	-	-	-	29.86	36.88	33.82	16.22	31.26	26.93	31.36
	Enh.	-	-	-	87.79	94.18	94.04	90.93	89.52	67.10	105.24
Pixz	-E	22.78	15.12	5.53	21.15	15.34	11.22	7.80	16.94	15.39	5.97
	+E	-	-	-	17.84	21.06	17.35	7.10	13.01	10.34	10.55
	Enh.	-	-	-	84.34	137.30	154.56	91.07	76.77	67.16	85.91
Gzip	-E	22.97	16.87	11.28	23.50	19.51	17.07	9.35	22.38	24.00	8.68
	+E	-	-	-	19.31	22.39	20.28	8.27	18.94	16.19	17.58
	Enh.	-	-	-	82.14	114.77	118.78	88.46	84.65	67.45	98.44
Bzip2	-E	18.77	12.46	7.98	19.49	14.47	13.19	5.33	15.92	15.87	6.63
	+E	-	-	-	16.01	16.85	13.56	4.87	12.13	10.39	11.63
	Enh.	-	-	-	82.15	116.48	102.78	91.39	76.21	65.48	85.22
7zip	-E	23.69	15.71	6.23	21.37	17.00	11.61	8.76	17.02	16.16	6.37
	+E	-	-	-	18.37	22.93	17.80	7.60	13.16	10.78	11.00
	Enh.	-	-	-	85.93	134.88	153.28	86.79	77.29	66.68	83.96
	Best $R_{AMIE \cdot i}$	**-**	**-**	**-**	**16.01**	**16.85**	**13.56**	**4.87**	**12.13**	**10.34**	**10.55**
	Best $E_{AMIE/i}$	**-**	**-**	**-**	**82.14**	**92.25**	**91.09**	**84.69**	**76.21**	**65.48**	**83.96**
	Best $E^{-1}_{AMIE/i}$	**-**	**-**	**-**	**19.62**	**19.20**	**16.26**	**7.41**	**22.60**	**17.59**	**18.20**
	Best Enhancement Efficiency	**-**	**-**	**-**	**99.34**	**95.16**	**91.51**	**77.67**	**70.45**	**89.76**	**69.06**

TABLE IV
ENHANCEMENT EFFECT - GZIP

	Dataset	E	D	S	DBf	Fs	Fm	U	FB	WN	N
	Reduc. Ratio (%)	22.97	16.87	11.28	23.50	19.51	17.07	9.35	22.38	24.00	17.86
LZ4	-E	41.04	31.32	19.42	36.60	47.37	43.61	18.66	33.42	38.86	14.08
	+E	23.29	17.05	10.68	23.59	15.94	8.96	9.36	22.38	24.00	17.86
	Enh.	56.74	54.42	55.01	64.44	33.65	20.55	50.15	66.96	61.75	61.63
Lzop	-E	36.13	27.05	17.83	34.01	39.16	35.97	17.84	34.92	40.13	14.48
	+E	24.05	17.48	11.32	23.84	20.69	10.24	9.37	22.38	24.00	17.86
	Enh.	66.58	64.63	63.50	70.11	52.83	28.47	52.55	64.10	59.80	59.94
Pixz	-E	22.78	15.12	5.53	21.15	15.34	11.22	7.80	16.94	15.39	5.97
	+E	23.96	17.42	10.37	23.77	16.25	8.89	9.37	22.38	24.00	17.86
	Enh.	105.17	115.18	187.46	112.37	105.95	79.25	120.16	132.09	155.94	145.39
Gzip	-E	22.97	16.87	11.28	23.50	19.51	17.07	9.35	22.38	24.00	8.68
	+E	23.35	17.08	10.36	23.61	16.27	8.95	9.36	22.38	24.00	17.86
	Enh.	101.64	101.24	91.90	100.43	83.42	52.44	100.07	100.02	100.02	100.02
Bzip2	-E	18.77	12.46	7.98	19.49	14.47	13.19	5.33	15.92	15.87	6.63
	+E	29.32	20.36	10.80	25.75	21.68	11.86	9.50	22.48	24.12	17.94
	Enh.	156.23	163.38	135.31	132.13	149.86	89.92	178.08	141.21	151.94	131.46
7zip	-E	23.69	15.71	6.23	21.37	17.00	11.61	8.76	17.02	16.16	6.37
	+E	25.17	18.09	10.42	24.19	18.10	9.33	9.39	22.38	24.00	17.86
	Enh.	106.22	115.20	167.26	113.20	106.46	80.34	107.25	131.48	148.49	136.32
	Best $R_{Gzip \cdot i}$	**23.29**	**17.05**	**10.36**	**23.59**	**15.94**	**8.89**	**9.36**	**22.38**	**24.00**	**17.86**
	Best $E_{Gzip/i}$	**56.74**	**54.42**	**55.01**	**64.44**	**33.65**	**20.55**	**50.15**	**64.10**	**59.80**	**59.94**
	Best $E^{-1}_{Gzip/i}$	**101.35**	**101.03**	**91.90**	**100.35**	**81.72**	**52.09**	**100.06**	**100.00**	**100.00**	**100.00**
	Best Enhancement Efficiency	**40.49**	**31.00**	**20.50**	**36.48**	**57.96**	**83.07**	**18.65**	**34.91**	**40.13**	**29.79**