A NOTE ON PROOFS OF THE PIGEONHOLE PRINCIPLE IN DEEP INFERENCE

ANUPAM DAS

ABSTRACT. It is known that the functional and onto variants of the propositional pigeonhole principle have polynomial-size proofs in the weakest deep inference systems. The unrestricted variant has quasipolynomial-size proofs when dag-like behaviour is permitted. Here we match that upper bound in systems free of dagness, and indeed other compression mechanisms, utilising a similar strategy to that of Atserias, Galesi and Gavalda for the monotone sequent calculus.

Preliminaries. We assume general familiarity with deep inference. As introductory material consult [BG09] and [Das12b] for known results on the complexity of deep inference, [Jeř09] for the relationship between deep inference and the monotone sequent calculus, [AGG00] for monotone sequent proofs of the pigeonhole principle and basic properties of threshold formulae, and [GG08] for an introduction to streamlining and atomic flows.

The proofs we ultimately obtain are free of compression mechanisms, or *uncompressed*, in the sense of [Das12a]. All proofs considered are monotone, i.e. there is no use of the negation rules, unless mentioned otherwise.

Systems are construed with constants $\bot, \top$ in the standard way, and we use the notation KS^+ to denote the system $\mathsf{KS} \cup \{\mathsf{aw}\uparrow, \mathsf{ac}\uparrow\}$. Bold lowercase letters $\boldsymbol{a}, \boldsymbol{b}, \ldots$ are used to denote vectors of propositional variables and bold uppercase letters $\boldsymbol{A}, \boldsymbol{B}, \ldots$ are used to denote matrices of propositional variables.

Threshold formulae and permutations. Threshold formulae assert that at least k out of n inputs are true. Logically, the truth of these formulae is preserved under permutation of the inputs, but it is not an easy task to provide propositional proofs of this fact in the absence of dag-like behaviour, provided in deep inference by the presence of *contraction loops*.

In this section we show that derivations with sequences of contraction loops of length $\log^{O(1)} n$ can be constructed for certain permutations: interleavings and thereby switching of rows and columns. This ensures that, when eliminating these loops, the derivations only blow up quasipolynomially.

Remark 1. Throughout, we will assume that m, n are powers of 2 and $m \leq n$.

Definition 2 (Threshold formulae)**.** We define monotone divide-and-conquer threshold formulae inductively as follows:

$$\mathsf{TH}^1_k(a) \equiv \begin{cases} \top & k = 0 \\ a & k = 1 \\ \bot & k > 1 \end{cases} \quad , \quad \mathsf{TH}^{2n}_k(\boldsymbol{a}, \boldsymbol{b}) \equiv \bigvee_{i+j=k} \mathsf{TH}^n_i(\boldsymbol{a}) \wedge \mathsf{TH}^n_j(\boldsymbol{b})$$

Observation 3. *$\mathsf{TH}^n_k(\boldsymbol{a})$ has size $n^{O(\log n)}$ and depth $O(\log n)$.*

Date: June 4, 2012.

Definition 4 (Interleaving)**.** For vectors $\boldsymbol{a} = (a_1, \dots, a_n), \boldsymbol{b} = (b_1, \dots, b_n)$ let $\boldsymbol{a} \between \boldsymbol{b}$ denote the *interleaving* of $\boldsymbol{a}$ with $\boldsymbol{b}$: $(a_1, b_1, \dots, a_n, b_n)$.

More generally, $(a_1, \dots, a_m, b_1, \dots b_m, \cdots, a_{n-m+1}, \dots, a_n, b_{n-m+1}, \dots, b_n)$ is the m-interleaving of $\boldsymbol{a}$ with $\boldsymbol{b}$, denoted $\boldsymbol{a} \between_m \boldsymbol{b}$.

Definition 5 (Distributivity)**.** We define distributivity rules as abbreviations for the following derivations:

$$\mathsf{dist}\uparrow\colon \quad \dfrac{\dfrac{\dfrac{A}{A \wedge A}\,\mathsf{c}\uparrow}{(A \wedge B) \vee (A \wedge C)}\,2{\cdot}\mathsf{s} \wedge [B \vee C]}{} \quad , \quad \mathsf{dist}\downarrow\colon \quad \dfrac{\dfrac{(A \wedge B) \vee (A \wedge C)}{A \vee A}\,\mathsf{m}}{A}\,\mathsf{c}\downarrow \wedge [B \vee C]$$

Definition 6 (Contraction loops)**.** See [Das12b].

Lemma 7. *There is a derivation* $\begin{array}{c}\mathsf{TH}^{2n}_k(\boldsymbol{a}, \boldsymbol{b}) \\ \Vert\,\mathsf{ks}^+ \\ \mathsf{TH}^{2n}_k(\boldsymbol{a} \between_m \boldsymbol{b})\end{array}$ *where each sequence of contraction loops has length* $O(\log n - \log m)$ *and width* $O(n)$.

Proof. We give an inductive step from n to $2n$. Let $\boldsymbol{a} = (a_1, \dots, a_n), \boldsymbol{b} = (b_1, \dots, b_n), \boldsymbol{c} = (c_1, \dots, c_n), \boldsymbol{d} = (d_1, \dots, d_n)$. The derivation is as follows:

$$= \dfrac{\mathsf{TH}^{4n}_r(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})}{\displaystyle\bigvee_{s+t=r} \left(\begin{array}{c} = \dfrac{\mathsf{TH}^{2n}_s(\boldsymbol{a}, \boldsymbol{b})}{\displaystyle\bigvee_{i+j=s} \mathsf{TH}^n_i(\boldsymbol{a}) \wedge \mathsf{TH}^n_j(\boldsymbol{b})} \wedge = \dfrac{\mathsf{TH}^{2n}_t(\boldsymbol{c}, \boldsymbol{d})}{\displaystyle\bigvee_{k+l=t} \mathsf{TH}^n_k(\boldsymbol{c}) \wedge \mathsf{TH}^n_l(\boldsymbol{d})} \\ \Vert\,\mathsf{dist}\uparrow \\ \displaystyle\bigvee_{\substack{i+j=s \\ k+l=t}} \big(\mathsf{TH}^n_i(\boldsymbol{a}) \wedge \mathsf{TH}^n_j(\boldsymbol{b})\big) \wedge \big(\mathsf{TH}^n_k(\boldsymbol{c}) \wedge \mathsf{TH}^n_l(\boldsymbol{d})\big) \end{array} \right)}$$

$$= \dfrac{\displaystyle\bigvee_{\substack{i+k=s' \\ j+l=t'}} (\mathsf{TH}^n_i(\boldsymbol{a}) \wedge \mathsf{TH}^n_k(\boldsymbol{c})) \wedge (\mathsf{TH}^n_j(\boldsymbol{b}) \wedge \mathsf{TH}^n_l(\boldsymbol{d}))}{\displaystyle\bigvee_{s'+t'=r} \left(\begin{array}{c} \Vert\,\mathsf{dist}\downarrow \\ \mathsf{TH}^{2n}_{s'}(\boldsymbol{a}, \boldsymbol{c}) \quad \mathsf{TH}^{2n}_{t'}(\boldsymbol{b}, \boldsymbol{d}) \\ IH\Vert \qquad \wedge \qquad IH\Vert \\ \mathsf{TH}^{2n}_{s'}(\boldsymbol{a} \between_m \boldsymbol{c}) \quad \mathsf{TH}^{2n}_{t'}(\boldsymbol{b} \between_m \boldsymbol{d}) \end{array} \right)}$$

$$= \dfrac{}{\mathsf{TH}^{4n}_r((\boldsymbol{a}, \boldsymbol{b}) \between_m (\boldsymbol{c}, \boldsymbol{d}))}$$

where derivations marked IH are obtained by the inductive hypothesis.

Each inductive step adds $O(\log n)$ contraction loops of width $O(n)$ in parallel, in the form of the distributivity steps, on top of two copies of the inductive hypothesis in parallel. The induction terminates in $O(\log \frac{n}{m})$ steps, whence the upper bound is obtained. □

Remark 8. When $m = 1$ we get the standard interleaving, with sequences of contraction loops of length $O(\log n)$ and width $O(n)$.

Convention 9. We consider a matrix to be equivalent to the vector obtained by a rows-first reading of it. E.g. $\boldsymbol{A} = (a_{ij})$ is considered to be equivalent to the vector $(a_{11}, a_{21}, \dots a_{n1}, a_{12}, \dots)$. In this way the *transpose* of the matrix, $\boldsymbol{A}^\intercal$, is equivalent to the vector obtained by a columns-first reading of $\boldsymbol{A}$.

Observation 10. *For matrices* $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *of equal dimensions,* $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}^\intercal = \begin{pmatrix} \boldsymbol{A}^\intercal & \boldsymbol{C}^\intercal \\ \boldsymbol{B}^\intercal & \boldsymbol{D}^\intercal \end{pmatrix}$.
(Of course, in any such situation, $\boldsymbol{A}$ *and* $\boldsymbol{D}$ *will also have equal dimensions).*

Theorem 11 (Transposition). *There is a derivation* $\begin{array}{l}\mathsf{TH}^n_k(\boldsymbol{X})\\ \|\mathsf{KS}^+\\ \mathsf{TH}^n_k(\boldsymbol{X}^\intercal)\end{array}$ *where each sequence of contraction loops has length* $O(\log^2 n)$ *and width* $O(n)$.

Proof. We give an inductive step from n to $2n$. Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$ be the four quadrants of $\boldsymbol{X}$. The derivation is as follows:

$$
\begin{array}{l}
\\
= \\
\\
= \\
\\
\\
\end{array}
\begin{array}{c}
\mathsf{TH}^{2n}_k \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \\
\hline
\bigvee_{i+j=k} \begin{pmatrix} \mathsf{TH}^n_i \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \end{pmatrix} \\ \|IH \\ \mathsf{TH}^n_i \begin{pmatrix} \boldsymbol{A}^\intercal \\ \boldsymbol{B}^\intercal \end{pmatrix} \end{pmatrix} \wedge \begin{pmatrix} \mathsf{TH}^n_i \begin{pmatrix} \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \\ \|IH \\ \mathsf{TH}^n_i \begin{pmatrix} \boldsymbol{C}^\intercal \\ \boldsymbol{D}^\intercal \end{pmatrix} \end{pmatrix} \\
\hline
\mathsf{TH}^{2n}_k \left(\begin{pmatrix} \boldsymbol{A}^\intercal \\ \boldsymbol{B}^\intercal \end{pmatrix}, \begin{pmatrix} \boldsymbol{C}^\intercal \\ \boldsymbol{D}^\intercal \end{pmatrix} \right) \\
\|\text{interleave} \\
\mathsf{TH}^{2n}_k \begin{pmatrix} \boldsymbol{A}^\intercal & \boldsymbol{C}^\intercal \\ \boldsymbol{B}^\intercal & \boldsymbol{D}^\intercal \end{pmatrix}
\end{array}
$$

where the derivations marked IH are obtained by the inductive hypothesis and Obs. 10, and the derivation marked *interleave* is obtained by applying Lemma 7 to interleave the rows of the two matrices.

Each inductive step adds an interleaving below two copies of the inductive hypothesis in parallel, thereby adding $O(\log n)$ contraction loops of width $O(n)$ in sequence by Lemma 7. The induction terminates in $O(\log n)$ steps, whence the upper bound is obtained. □

The pigeonhole principle. The propositional pigeonhole principle asserts that if $n-1$ holes contain n pigeons then there are at least two pigeons in the same hole. In its unrestricted formulation pigeons are permitted to be in multiple holes.

Definition 12 (Pigeonhole principle)**.** We define the following:

$$\mathsf{LPHP}_n \equiv \bigwedge_{i=1}^{n} \bigvee_{j=1}^{n-1} a_{ij} \quad \mathsf{RPHP}_n \equiv \bigvee_{j=1}^{n-1} \bigvee_{i=1}^{n} \bigvee_{i'=i+1}^{n} (a_{i'j} \wedge a_{ij}) \quad \mathsf{PHP}_n \equiv \mathsf{LPHP}_n \to \mathsf{RPHP}_n$$

Definition 13. Let $\bot_{mn}$ be the $(m \times n)$ matrix with the constant $\bot$ at every entry. Define $\boldsymbol{A}_n = \begin{pmatrix} (a_{ij}) & \bot_{n1} \end{pmatrix}$, with i, j ranging as in Dfn. 12. I.e. $\boldsymbol{A}_n$ is obtained by extending (a_{ij}) with an extra column of $\bot$-entries, so that it is a square matrix.

Proposition 14. *There are derivations* $\begin{array}{l}\mathsf{LPHP}_n\\ \|\{\mathsf{w}\downarrow,\mathsf{c}\uparrow\}\\ \mathsf{TH}^{n^2}_n(\boldsymbol{A}_n)\end{array}$ *and* $\begin{array}{l}\mathsf{TH}^{n^2}_n(\boldsymbol{A}^\intercal_n)\\ \|\{\mathsf{w}\uparrow,\mathsf{c}\downarrow\}\\ \mathsf{RPHP}_n\end{array}$.

Proof. A simple exercise, but essentially contained in [AGG00] under the translation from monotone sequent calculus to deep inference in [Jeř09]. □

Definition 15. For a system $\mathcal{S}$, let $\overline{\mathcal{S}}$ denotes the set of its dual rules. Derivations of the form $\begin{array}{l}A\\ \|\overline{\mathsf{KS}^+}\\ B\\ \|\mathsf{KS}^+\\ C\end{array}$ and $\begin{array}{l}A\\ \|\overline{\mathsf{KS}}\\ B\\ \|\mathsf{KS}\\ C\end{array}$ are called *weakly streamlined* and *strongly streamlined* resp.

Remark 16. The sets of (strongly) streamlined derivations defined above are actually strict subsets of those given in [GG08] and [GGS10], where more liberal geometric definitions are given via atomic flows, but are essentially equivalent via polynomial transformations.

Note also that, if A, B are monotone formulae, then any (strongly) streamlined derivation $\begin{array}{l} A \\ \| \\ B \end{array}$ is negation-free and so contains no negation rules.

Lemma 17. *A weakly streamlined derivation of size S and whose sequences of contraction loops have length at most l and width at most w can be transformed to a strongly streamlined derivation with same premiss and conclusion and size $w^l \cdot S$.*

Proof. See [Das12b]. □

Theorem 18. *There are strongly streamlined derivations* $\begin{array}{l} \mathsf{LPHP}_n \\ \| \\ \mathsf{RPHP}_n \end{array}$ *of size $n^{O(\log^2 n)}$.*

Proof. Prop. 14 and Thm. 11 yield derivations with the required premiss and conclusion and with sequences of contraction loops of length $O(\log^2 n)$ and width $O(n)$. The result then follows by Lemma 17. □

Proposition 19. *A strongly streamlined derivation from A to B can be linearly transformed to a* KS *proof of $\bar{A} \vee B$.*

Proof. Choose a formula in the 'middle' of the derivation, i.e. a formula C such that the derivation is of the form $\begin{array}{l} A \\ \Phi \| \overline{\mathsf{KS}} \\ C \\ \Psi \| \mathsf{KS} \\ B \end{array}$. Now construct the KS proof $\mathsf{i}{\downarrow}\dfrac{\top}{\bar{C} \quad C}$ $\bar{\Phi} \| \mathsf{KS} \vee \Psi \| \mathsf{KS}$ $\bar{A} \quad B$. □

Corollary 20. *There are* KS *proofs of* PHP_n *of quasipolynomial size.*

Remark 21. Though we have assumed that n is a power of 2, the proof is actually sufficient for all n, as pointed out in [AGG00], by just mimicking the proof for the next power of 2 and replacing some variables by $\bot$.

Ongoing and future work. We have constructed quasipolynomial-size strongly streamlined derivations that interleave and transpose inputs of threshold formulae; doing the same for all permutations is the subject of ongoing work. Weakly streamlined derivations for arbitrary permutations are essentially givin in [AGG00], under the translation in [Jeř09]. The problem with the present method is that interleavings do not form a generating set for the symmetric group. However a generalisation of them, corresponding to the set of possible card shuffles on a deck of size n, do form such a set, and even one that can reach any permutation in $O(\log n)$ steps, as required. The task at hand is now solely to formalise this procedure in deep inference proofs.

Note that the proofs we have given, albeit $n^{O(\log^2 n)}$ so quasipolynomial in size, are not in polynomial correspondence with those constructed in [AGG00] for the monotone sequent calculus, and essentially KS^+, which are of a smaller quasipolynomial size, $n^{O(\log n)}$. In fact it is conjectured that there are polynomial-size proofs in KS^+, due to the more general conjecture that the monotone sequent calculus polynomially simulates the full sequent calculus over monotone sequents, where polynomial-size proofs exist [Bus87]. Finding more efficient proofs in either KS or KS^+ would be an interesting future pursuit.

REFERENCES

[AGG00] Albert Atserias, Nicola Galesi, and Ricard Gavalda. Monotone proofs of the pigeon hole principle. 2000.

[BG09] Paola Bruscoli and Alessio Guglielmi. On the proof complexity of deep inference. *ACM Transactions on Computational Logic*, 10(2):1–34, 2009. Article 14. http://cs.bath.ac.uk/ag/p/PrComplDI.pdf.

[Bus87] Samuel R. Buss. Polynomial size proofs of the propositional pigeonhole principle. *Journal of Symbolic Logic*, 52(4):916–927, 1987.

[Das12a] Anupam Das. Characterising aspects of proof compression. http://people.bath.ac.uk/ad402/items/CompMech/CompMech.pdf, 2012.

[Das12b] Anupam Das. Complexity of deep inference via atomic flows. http://people.bath.ac.uk/ad402/items/RelComp/RelComp.pdf, 2012.

[GG08] Alessio Guglielmi and Tom Gundersen. Normalisation control in deep inference via atomic flows II. http://cs.bath.ac.uk/ag/p/NormContrDIAtFl2.pdf, 2008.

[GGS10] Alessio Guglielmi, Tom Gundersen, and Lutz Straßburger. Breaking paths in atomic flows for classical logic. In Jean-Pierre Jouannaud, editor, *25th Annual IEEE Symposium on Logic in Computer Science*, pages 284–293. IEEE, 2010. http://www.lix.polytechnique.fr/~lutz/papers/AFII.pdf.

[Jeř09] Emil Jeřábek. Proof complexity of the cut-free calculus of structures. *Journal of Logic and Computation*, 19(2):323–339, 2009. http://www.math.cas.cz/~jerabek/papers/cos.pdf.