A Appendix

A.1 Data pre-processing

In this section, the various pre-processing steps of data performed at different stages are explained.

Extracting (document, summary) pairs: The 120 pairs of Amicus Briefs were scrapped from their website[11]. The Summary of Arguments section of the Amicus Briefs was extracted as the target summary and the main content excluding title page, table of contents, acknowledgements, appendix etc was extracted as document/source.

Sentence pre-processing: Sentences from the (document, summary) files were split using the spaCy[12] sentence splitter. Furthermore, the sentences were each processed to remove special characters using regex rules. If a sentence contained less that 5 words, it was pruned out from the computation of $f(s, t)$ to reduce the complexity of pairs considered.

A.2 Sentence Saliency Classifier

Training Details: Our classifier uses the BERT sequence labeling configuration[13] from `transformers` (Wolf et al., 2019), which is a pretrained BERT-base model with an initially untrained classification head on the `[CLS]` feature vector. This model is then finetuned for 5 epochs using the training data which consists of 5363 sentences in the Amicus dataset (equal distribution among the two classes). We use a train / dev / test split of 60%, 20%, 20%. Training configuration of the classifier is as follows: learning rate = 2e-5, max_grad_norm = 1.0, num_training_steps = 1000, num_warmup_steps = 100, warmup_proportion = 0.1, Optimizer = Adam, Schduler = linear with warmup.

Alternate methods to choose +/- samples: The aggregate scoring method mentioned in Section 2 was one choice to pick salient and non-salient samples for each document. Aggregate method compresses the source by 61% on an average. The other methods experimented were:

- Top k - Bottom k: $\forall t_j \in$ T, we picked the top-k scoring source sentences as positive samples and the bottom-k sentences as the negative samples ensuring that $\{\text{positive}\} \cap \{\text{negative}\} = \emptyset$. Using this technique, the classifier achieves accuracy of nearly 1 as can be seen from Table 5. On qualitative analysis, we identified that there is a clear distinction in the positive and the negative examples. Eg: sentences such as 'This document is prepared by XYZ' would be picked as non salient sentence and classifier is able to achieve high accuracy. This could however be used to train a classifier to *identify boiler plate sentences* across the document. This method compresses source document by 63% on an average.

- Random negative sampling: Salient examples were chosen for a document as per the above method. For the non salient examples, we randomly sampled from the rest of the document. This allows the classifier to learn about sentences that that are difficult to be classified as positive or negative. Hence, the accuracy of the classifier is lower than the other two methods as can be seen from Table 5. This method compresses the source document by 70% on an average.

Compute time and resources: Execution time for different choice of *f(s,t)* for all 120 pairs:

- Perplexity using GPT-2:executes within 15hrs using 2 GPUs

- Entailment score using RoBERTa: executes within 22hrs using 2 GPUs

- Cosine Similarity using BERT [CLS] embeddings: executes within 3hrs on a single GPU

- BLEU score using nltk: executes within 15min on a single GPU.

These scores need to be generated once and can be reused for various experiments. Sampling methods to choose salient and non-salient sentences for each document takes less than a minute to run.

Analysis: (a) Table 5 shows the classifier accuracies for combinations of *f(s,t)* and sampling methods. We observe that for the aggregate sampling method, although perplexity based classifier does not have the highest accuracy, our

[11]https://publichealthlawcenter.org/amicus-briefs

[12]https://pypi.org/project/spacy/

[13]`https://huggingface.co/transformers/model_doc/bert.html#transformers.BertForSequenceClassification`

Sampling Method	f(s,t)	Accuracy
Aggregate scoring for each source sentence.	BLEU	0.7813
	Perplexity	0.7366
	Entailment	0.6569
	Similarity	0.8391
Top k-Bottom k sources sentences or each summary sentence	BLEU	0.9997
	Perplexity	0.9915
	Entailment	0.9973
	Similarity	1
Top k for each summary sentence and random negative sampling from the remaining document.	BLEU	0.5784
	Perplexity	0.655
	Entailment	0.5611
	Similarity	0.6233

Table 5: The accuracy of the held out set of Amicus for different classifiers trained on the data prepared using choice of different f(s,t) and sampling methods. Here, k=3.

pipeline where $f(s,t)$ is perplexity score gives the best result(ROUGE) amongst the ablation experiments(Table 4). Classifier accuracy is determined on automated labelling based on the saliency score, rather than true labels, hence best classifier does not imply best summarization. (b) Table 6 shows the examples of using perplexity as *f(s,t)* to see how the summary grounds the source. The table shows three summary sentences and the corresponding source sentences that had the lowest perplexity scores. We can see that, summary either has a similar meaning or logically follows the source. (c) Table 7 has three examples each for salient sentences and non-salient sentences inferred by the classifier trained on data prepared as mentioned in Section 2. The third sentence in the non-salient sentences column is an example of boiler-plate content detected that is present across documents.

A.3 Abstractive Summarizer: BART

BART is a seq2seq model based on denoising pre-training objective which is supposed to generalize better on various natural language understanding tasks; abstractive summarization being one of them. For abstractive stage of our proposed approach, we decided to see (*bart.large.cnn*) variant which is essentially BART-large model (with 12 encoder and decoder layers and 400 million parameters) finetuned for CNN/DM summarization task. We use the pre-computed weights available for use here[14]. Using BART's text generation script, we set length penalty (lenpen) as 2.0 and minimum length (min_len) as 500 words in order to encourage BART to produce longer outputs which is more suitable to our dataset. Also, we use beam size of 4 and and no_repeat_ngram_size of 3.

Finetuning: We use the train and dev splits of Amicus dataset (96 source-target pairs) and finetune BART for summarization task starting from its CNN/DM finetuned checkpoint. First, we pre-process the dataset as per the guidelines in the official code[15]. We finetune for 500 epochs with learning rate of 3e-5 and early stop if validation loss doesn't decrease for 50 epochs. Others parameters are as follows: total_num_updates = 20000, warmup_updates = 500, update_freq = 4, optimiser = Adam with weight decay of 0.01. Rest of parameters were kept as default in the official script. Results (Precision, Recall, F1) on the test set of Amicus using the existing BART model and finetuned BART are shown in Table 8.

Table 9 shows an example of target summary and summary generated by our model(Section 2) for one sample source document. We can see that the summary generated by our model is fluent and has coherent flow of information.

[14] https://github.com/pytorch/fairseq/tree/master/examples/bart

[15] https://github.com/pytorch/fairseq/blob/master/examples/bart/README.summarization.md

Summary Sentence	Source Sentence
In the immigration context, this jurisprudence has prompted the Court to reject the notion that the so-called entry fiction is of constitutional significance.	Prior to Knauff and Mezei, the distinction between noncitizens who had entered the United States and those who remained outside it had not had been elevated to a bright-line constitutional rule, and entry had never been completely determinative of the fact or extent of protection under the Due Process Clause.
It has accordingly authorized such detention only in limited circumstances pursuant to a carefully defined scheme.	The Court's substantive due process jurisprudence also recognizes that an individual may be subjected to regulatory detention only in narrow circumstances under a carefully drawn scheme.
With respect to substantive due process, this Court has increasingly recognized the punitive consequences of indefinite regulatory detention.	Thus, the Court has substantially restricted the availability and duration of regulatory confinement in the — years since it decided Meze1.In Zadvydas, this Court established that its substantive due process jurisprudence provided the appropriate framework for evaluating the administrative detention of noncitizens pending removal from the United States.

Table 6: Using GPT-2 perplexity as f(s,t), here are three sentences from the summary with corresponding source sentence, having the lowest perplexity.

Salient Sentences	Non-Salient sentences
The same time, the Court has long been skeptical of the military's authority to try individuals other than active service personnel.	A government predicated on checks and balances serves not only to make Government accountable but also to secure individual liberty.
On the basis of this revised test, the Court of Appeals refused to apply the exceptional circumstances exception to Al-Nashiri's petition.	At present, the Rules for Courts-Martial require that the accused be brought to trial within 120 days after the earlier of preferral of charges or confinement.
Consonant with that tradition, this Court should review the Court of Appeals' decision to confirm that exceptional delay before trial remains of central concern on habeas review and is indeed one of the very dangers the writ of habeas corpus was designed to avoid.	Respectfully submitted, May 31, 2017 LINDA A. KLEIN Counsel of Record AMERICAN BAR ASSOCIATION 321 North Clark Street Chicago ...

Table 7: This table shows the sentences classified as salient and non-salient from one Amicus source document. We can see that the last sentence in the non-salient sentences column shows an example of boiler-plate content present across documents. The classifier is trained on data chosen on aggregate score of source sentences where *f(s,t)* is GPT-2 perplexity.

Metric		BART	Ours + BART	f.t. BART	Ours + f.t. BART
ROUGE-1	Recall	40.87	47.46	46.90	56.04
	Precision	47.21	49.97	48.68	46.16
	F-1	40.17	44.97	43.47	47.07
ROUGE-2	Recall	13.76	16.54	17.84	21.50
	Precision	15.46	17.04	17.84	17.10
	F-1	13.36	15.37	16.30	17.64
ROUGE-L	Recall	18.34	25.58	21.30	29.62
	Precision	21.04	26.27	21.35	23.47
	F-1	17.95	23.95	19.35	24.40

Table 8: Overall pipeline results by adding our extractor (*f(s,t)* as GPT-2 perplexity + Classifier) to BART and finetuned BART (f.t. BART), including the precision and recall values for each metric.

This Court's determination of whether due process under the New HampshireConstitution requires court-appointed counsel for indigent parent-defendants, in order to protect their fundamental right to parent, requires the balancing of three factors–(1) theprivate interest at stake, (2) the risk of error and the value of procedural safeguards, and (3)the state's interest. See In re Shelby R., 148 N.H. 237, 240 (2002) (citing In re Richard A., 146 N.H..295, 298 (2001)). Because there is no dispute that the fundamental right to parent isat stake in abuse and neglect proceedings, the ABA focuses its discussion on the second and third factors of the three factor test.As to the second, so-called "risk of error" factor, the ABA's conclusion, after years of investigation and analysis, is that the absence of counsel for indigent parent-defendants in abuse and neglect proceedings results in a significant risk of an erroneous determination. This is especially true where the opposing party is the State. As to the third, state's interest factor, the ABA's investigation shows that the interests of both the parent and the state are best served where indigent parent-defendants are represented. The ABA respectfully suggests that the evidence and analysis relevant to these two factors is so compelling in most, if not all, abuse and neglect proceedings involving indigent parent-defendants, that a case-by-case balancing of the factors should be rejected in favor of a rule requiring the appointment of counsel] for indigent parent-defendants in all such proceedings. The evidence and analysis supporting the ABA's policy includes the fact that a substantial majority of states have recognized an unqualified right to counsel for indigentparent-defendants in child custody proceedings. Similarly, other industrial democraciesprovide indigent parent-defendants with such right to counsel. The ABA respectfully submits that this Court should require no less as a matter of due process under the New Hampshire Constitution.Although of whetherJn re Shelby R. resulted in a or not a natural parent'splurality role inruling, the the familyCourt is awas not split fundamentalon the libertyquestion interestprotected by the State Constitution. See In re Shelby R., 148 NH. at 244 (dissenting opinion).
Hampshire constitution requires this court to determine whether indigentparents have a legally protected interest. Most indigent parent - defendants are incapable of performingthe advocacy functions required in abuse and neglectproceedings. Most unrepresented parents cannot perform the advocacy functions - - including investigating facts , making an orderly factual presentation , and cross - examining witnesses - - that are required. The intense, emotionally charged backdrop against which custody decisionsare often made further exacerbates the inherent disadvantages faced by unrepresented indigent parents. The need for counsel for the indigentparent - defendant is especially great where the opposing party is the state. The court must weighthree factors : (1) the private interests that will be affected. (2) the risk of erroneousdeprivation of the liberty interest through the procedures used and the value , if any, ofadditional or substitute procedural safeguards. (3) the state ' s interest , including the function involved and fiscal and administrative burdens that additional or substituteprocedural requirements would entail id at 240 ; see also in re father , 155 n . h . 93 , 95 (2007) . this court has previously concluded as to the first factor that adversary child custody proceedings implicate a fundamental liberty interest - - the right to parent in this case, the central question thus becomes whether that right is sufficiently protected. The conclusion that counsel must be provided is so compelling in most , if not all cases , that a case - by - case balancing of the factors should be rejected in favor of a rule requiring the appointment of counsel for lowincome parent - defendant in all such proceedings to be constitutionally acceptable. The state is not the only adversary finding the only meaningful right to be heard when her adversary is not represented by counsel is not spaled against the traditional weapons of the state, such as the state's attorney general. The courts must also weigh the public interest in the child custody case, including the function involved and the cost of additional or substitute safeguards, as well as the cost to the state of the additional or substituted safeguards. The risk of an erroneous deprivation of the findamentalright to parent only increases the only increase in the risk that the state will find the child is not heard when the state is the adversary. The public interest is only increased by the fact that the child will not be heard by the state when the parent is represented by a lawyer. The high level of complexity of child custody cases makes it difficult for the court to make a fair and just decision.

Table 9: The table shows the comparison of summaries where the top summary is the target summary and the bottom summary is the one generated by our extractor and f.t BART. As we can see, the summary is coherent and has fluent information flow.