Mapping Regulatory System of WTO Dispute Settlement Body Using Deep Learning

Suyeol Yun

November 10, 2020

Abstract

1 Introduction

The Dispute Settlement Body (DSB) of the World Trade Organization (WTO) deals with trade disputes between WTO members. WTO members can file a lawsuit in the DSB to claim their impaired benifit related to the WTO agreements as a result of possible illegal action of the other member's trade policy. Then a judicial body, *Panel* or *Appellate Body* adjudicates the dispute and submits a report in which it expresses its conclusion as to whether the challenged trade policy is inconsistent to the rules of the WTO or not (World Trade Organization, 2017).

A lawsuit tends to cite multiple rules of the WTO agreement because one simple rule can't cover the complex characteristics of the trade policy that led to the dispute (Palmeter and Mavroidis, 2004). For example, the United States enacted *Continued Dumping and Subsidy Act of 2000* (CDSOA) that distributes the collected anti-dumping duties to its affected domestic producers. This act was challenged by other members with multiple rules of the WTO agreements such as the rules of *Anti-dumping* and *Subsidy* because this distribution could constitute an illegal subsidy as well as the breach of the rules of the anti-dumping (*See* Figure 1).

Citation of the rules of the WTO agreements tends to get more complicated because members cite the rules of the WTO agreements strategically. For example, members cite different rules of the WTO agreements to limit or to encourage the third party participation to the case. Since the third party participation can lead to early settlement of the dispute without continuous legal battle, members cite differntly according to their intention to settle the case earlier out of court or vice versa. (Johns and Pelc, 2014).

8.1 In the light of our findings, we conclude that **the CDSOA is inconsistent with AD (Anti-dumping) Articles 5.4, 18.1 and 18.4, SCM (Subsidy and Countervailing Measure)** Articles 11.4, 32.1 and 32.5, Articles VI:2 and VI:3 of the GATT 1994, and Article XVI:4 of the WTO Agreement.

...

8.3 **The CDSOA is a new and complex measure, applied in a complex legal environment**. In concluding that the CDSOA is in violation of the above mentioned provisions, we have been confronted by sensitive issues regarding the use of subsidies as trade remedies. this matter through negotiation.

Figure 1: **Excerpt from the Panel report for the *US - Offset (Byrd Amendment)* case:** Panel explicitly expresses the complexity of the trade policy (CDSOA) at issue and cites the rules of anti-dumpig (AD) and subsidy (SCM) at the same time to cover its complexity.

In addition to it, citation tends to get complicated if we consider the fact that *Panel* or *Appellate Body* defers to legal precedents. Legal precedents refer to its own judicial decisions and these precedents provide authoritative reference for deciding subsequent identical or similar cases. Members try to reshape these legal precedents in favor of their future interest rather than simply using the WTO DSB to resolve their trade dispute with other members (Pelc, 2014). For example, members tend to cite their favorable previous cases more often in specific issue areas where they face ligitagion more frequently with other members. (Strezhnev, 2014).

As WTO sets its main principles to regulate the world trade system, such as *Market Acceess* (across borders), *Non-discrimination* (between members or between domestic products and imported products) and *Transparency* (in publication and maintaining of each member's internal regulations), it's intellectually intriguing to understand how regulatory system of WTO DSB is structured to achieve these main principles (*See* Figure 2). By understanding this structure, we can improve WTO system to serve its main prinicples more effectively and to adopt to constantly changing world trade circumstances (Fredebeul-Krein and Freytag, 1999; Shaffer, 2004; Hufbauer and Stephenson, 2007).

However, it is extremely difficult to understand how rules of the WTO agreements are working together to achieve those main principles of WTO. This is because each citation is closely related to complex characteristics of each trade policy as exemplified in the above mentioned CDOSA case. Moreover, to understand interactions between multiple rules of WTO agreements properly, it requires one to generalize members' strategic citation patterns which are limited to each member's special interest rather than explaining the regulatory system of WTO DSB in general.

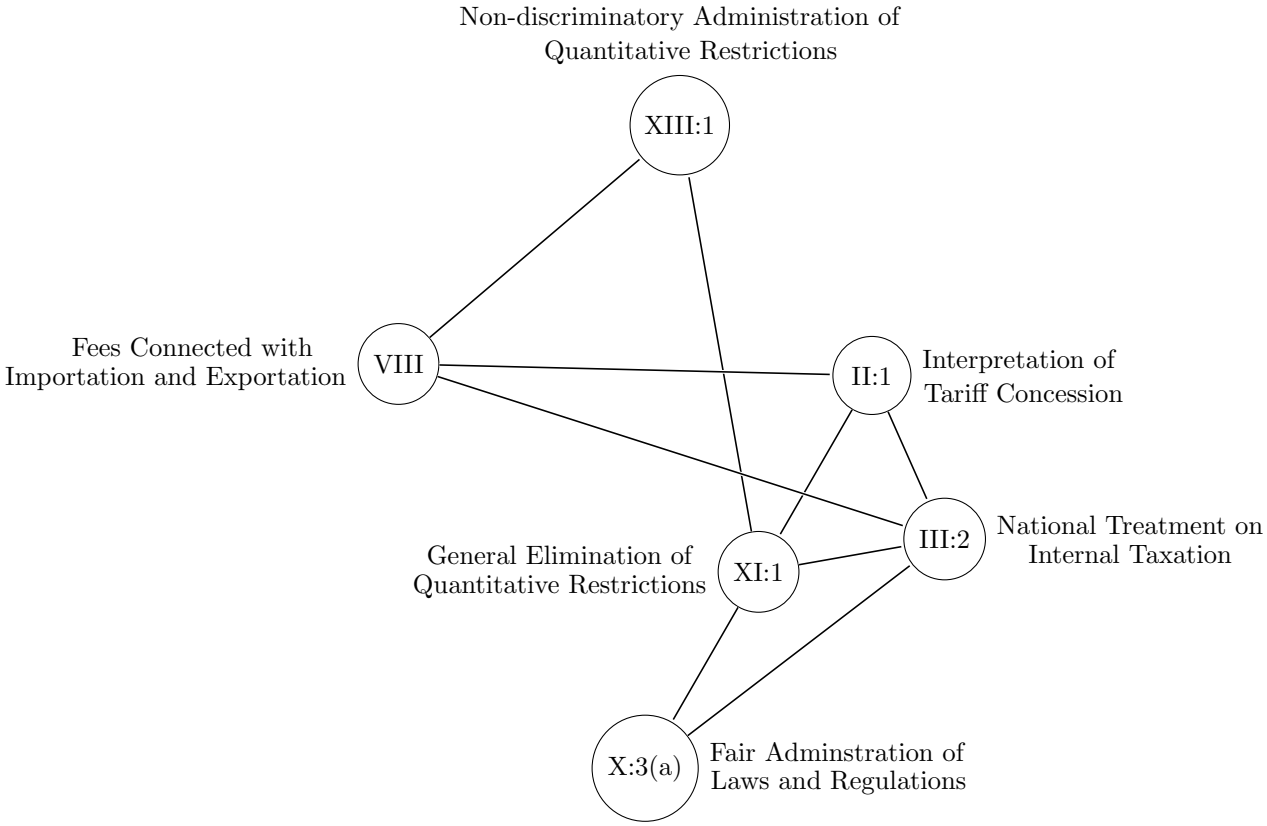


Figure 2: **Network of the Articles that Achieves *Market Access*:** This figure demonstrates a network of articles of WTO agreements that cooperatively achieves *Market access* principle of WTO. Tariff and Non-tariff barriers such as quantitative restriction, internal taxations and extra fees for crossing border can inhibit the chance of foreign goods to access the foreign market. Therefore, these articles tend to work together to ensure the *Market access* principle working properly. (Directions and weights of network edges are omitted for brevity.)

To address this issue, this paper maps the regulatory system of WTO DSB as a network of legal articles of the WTO agreements as formally defined in Figure 3 and illustrated in Figure 4. This is because the rules of the WTO agreements explicitly requires judicial bodies to address relevant articles together to construct its jurisprudence (*See* Figure 5). Upon this requirement, judicial bodies refer to multiple articles of the WTO agreements together to identify the complex legal identity of the trade policy that led to the dispute. In addition to it, judicial bodies cite multiple articles together to provide an authoritative interpretation of the rules of the WTO agreements (*See* Figure 4(b))

To map the regulatory system of WTO DSB as a network of legal articles of WTO agreement successfully, this paper designs a deep neural network (Figure **??**) that processes two different types of textual information. One is textual description of the dispute (*See* an example at Appendix A.1) and the other one is the textual content of a legal article of the WTO agreements (*See* an

Network of legal articles of WTO agreements is defined as

$$\textit{directed weighted graph } G = (V, E, w)$$

$$\text{where } V = \{v \mid v \text{ is a legal article of WTO agreement}\},$$

$$\vec{E} = \{(v_i, v_j) \mid (v_i, v_j) \in V \times V)\} \text{ and}$$

$$w : V \times V \to \mathbb{R}_+ \; s.t. \; w(v_i, v_i) = 0 \text{ and } \sum_{v_j \in V} w(v_i, v_j) = 1 \; \forall v_i \in V$$

Then define *edge weight matrix* $W = (w_{ij}) \in \mathbb{R}^{|V| \times |V|}$ *s.t.* $w_{ij} = w(v_i, v_j)$

Figure 3: **Formal Definition of Network of Legal Articles of WTO agreements:** I define network of legal articles of WTO agreements as a directed weighted graph where the sum of all weights coming out of a node sum up to 1 as illustrated in Figure 4

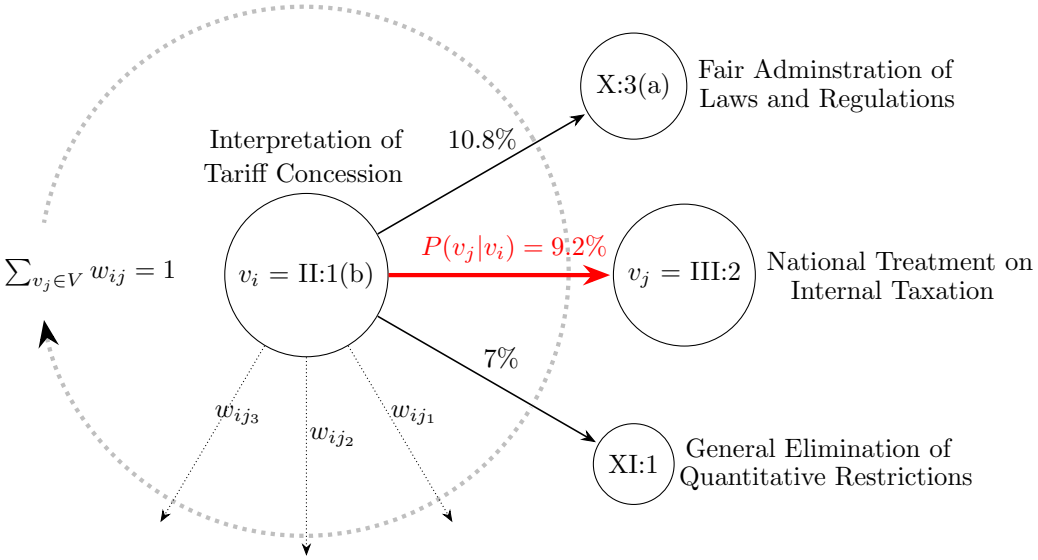
example at Figure 6). This design is improvised to mimic the reasoning process of WTO legal practitioners where the legal practitioners read the textual description of factual circumstances of the dispute and imagine regulatory contents of the applicable legal articles while he/she reads the factual description.

To train this neural network, this paper collected textual description of trade policy that led to the dispute and articles of the WTO agreement cited for each dispute case requested to the WTO DSB from 1995 to 2018 (Total 143 cases. *Check* the list in Appendix A.2). Using this collected data, I trained the neural network by enforcing the neural network to answer correctly whether a given article of the WTO agreements can be cited for the given textual description of trade policy that led to the dispute.

After finish training, I collected all the answers from the trained neural network (Figure **??**) and fitted a network of legal articles of the WTO agreement using this collection of answers. Since a network G comprises V (set of articles), E (directed edges) and their w (weights) as defined in Figure 3, I fitted a best set of weights w^* for given V and E using a machine learning technique called GENIE3 (Huynh-Thu et al., 2010) which is widely used in the biomedical engineering to reconstruct gene regulatory networks.[1]

To check whether this fitted network of WTO agreements G^* maps the regulatory system of WTO DSB properly, this paper compares the created network and the jurisprudence of WTO DSB made by *Panel* and *Appellate Body*. This comparison reveals that the fitted network G^*

[1] Anaology of international normative system to genetics maybe natural because gene expressions (achieving main principles of WTO) are governed by complex interaction between multiple regulatory proteins (interaction between legal articles of WTO). Similar notion is adopted in Florini (1996) to explain the the evolution of norm of transparency in international security.



(a) **Illustrated edge weights of a source node Article II:1(b)**

"The dictionary definition of the noun 'excess' is '[t]he amount by which one number or quantity exceeds another'. More specifically, 'in excess of' means 'more than'. Thus, as a textual matter, a particular number or quantity is 'in excess of' another number or quantity if it is greater, regardless of the extent to which it is greater. ***Looking at the context of Article II:1(b), first sentence, we note that Article III:2, first sentence, of the GATT 1994 is cast in very similar terms and in fact uses the phrase 'in excess of'***:

> The products of the territory of any contracting party imported into the territory of any other contracting party shall not be subject ... to internal taxes or other internal charges of any kind in excess of those applied ... to like domestic products ...

(b) **Exerpt from the panel report for the *Russia – Tariff Treatment* case:** Panel clarifies the meaning of the *'in excess of'* in Article III:2 with an anology to the Article II:1(b).

Figure 4: **Illustration of Network of Legal Articles of WTO agreements:** Every directed edge weight is interpreted as the conditional probability $P(v_j|v_i)$ of how probably a source node v_i constitutes a legal context to clarify the meaning of the target node v_j among all target nodes $v_{j \neq i} \in V$. Above subfigure (a) represents how jurisprudence of Panel stated in (b) is represented as an edge weight where source node Article II:1(b) constitutes the legal context of the target node Article III:2 with the probability of 9.2% among all possible target articles.

Article 7

Terms of Reference of Panels

1. Panels shall have the following terms of reference unless the parties to the dispute agree otherwise within 20 days from the establishment of the panel:

> "To examine, **in the light of the relevant provisions** in (name of the covered agreement(s) cited by the parties to the dispute), the matter referred to the DSB by (name of party) in document ... and to make such findings as will assist the DSB in making the recommendations or in giving the rulings provided for in that/those agreement(s)."

2. **Panels shall address the relevant provisions** in any covered agreement or agreements cited by the parties to the dispute. ...

Figure 5: **Article 7 of the Dispute Settlement Understanding (DSU):** DSU provides a legal guidelines on how judicial boides of WTO shall adjudicate the requested disputes. It explictly requires judicial bodies to interweave relevant articles of the WTO agreements to clarify it's meaning, scope and interpretation.

captures the interaction between the articles of WTO agreements similarly with the jurisprudence of *Panel* and *the Appellate Body*. This similarity guarantees that the fitted network G^* closely maps the regulatory system of WTO DSB since only these two judicial bodies can authoritatively consitute the jurisprudence over how rules of WTO agreements are working together to achieve the main principles of WTO.

Morover, this paper justifies the use of textual information inside the textual description of the dispute and legal article by showing that simply using the co-citation pattern between articles of the WTO disputes can't qualitatively fit the G^* (*See* Figure **??** and Figure **??**). Upon this necessity of using the textual information, this paper also justifies the use of neural network that is computationally intensive since it's generally known that proper design of neural network is able to effectively extract information from the textual content.

Finally, I will explain how this fitting of network of legal articles can contribute to the current study of international normative system.

2 Data: Types, Composition and Collection Process

This section explains the composition of data and its collection process in detail.

Article I

General Most-Favoured-Nation Treatment

1. With respect to customs duties and charges of any kind imposed on or in connection with importation or exportation or imposed on the international transfer of payments for imports or exports, and with respect to the method of levying such duties and charges, and with respect to all rules and formalities in connection with importation and exportation, and with respect to all matters referred to in paragraphs 2 and 4 of Article III, any advantage, favour, privilege or immunity granted by any contracting party to any product originating in or destined for any other country shall be accorded immediately and unconditionally to the like product originating in or destined for the territories of all other contracting parties...

Figure 6: **Example of content of a legal article of the WTO agreement:** Article I:1 of GATT 1994 that prohibits the discrimination among similar products (WTO DSB prefer to call *like products*).

2.1 Overview: How Members Raise Claims in WTO DSB

As explained in the introduction, a trade policy that led to a dispute (preferably called as *Government Measure* in WTO DSB) is pretty much complicated as explictly expressed by the Panel in Figure 1.

To address this complexity, members who raise the claim (preferably called *complainant* in WTO DSB) usually cite multiple articles of the WTO agreements at the same time. For example, in the *US - Offset (Byrd Amendment)* case, a group of complainants cited articles as shown in Table 1 from the WTO agreements to claim the inconsitency of *Continued Dumping and Subsidy Act of 2000* (CDSOA) of United States to these articles[2]:

Name of WTO Agreement	Cited Articles
Agreement on Anti-dumping	1, 5.4, 8, 18.1, 18.4
General Agreement on Tariffs and Trade 1994	VI:3, X:3, XXIII:1, VI:2
Agreement on Subsidies and Countervailing Measures	4.10, 7.9, 10, 11.4, 18, 32.1, 32.5
Agreement Establishing the World Trade Organization	XVI:4

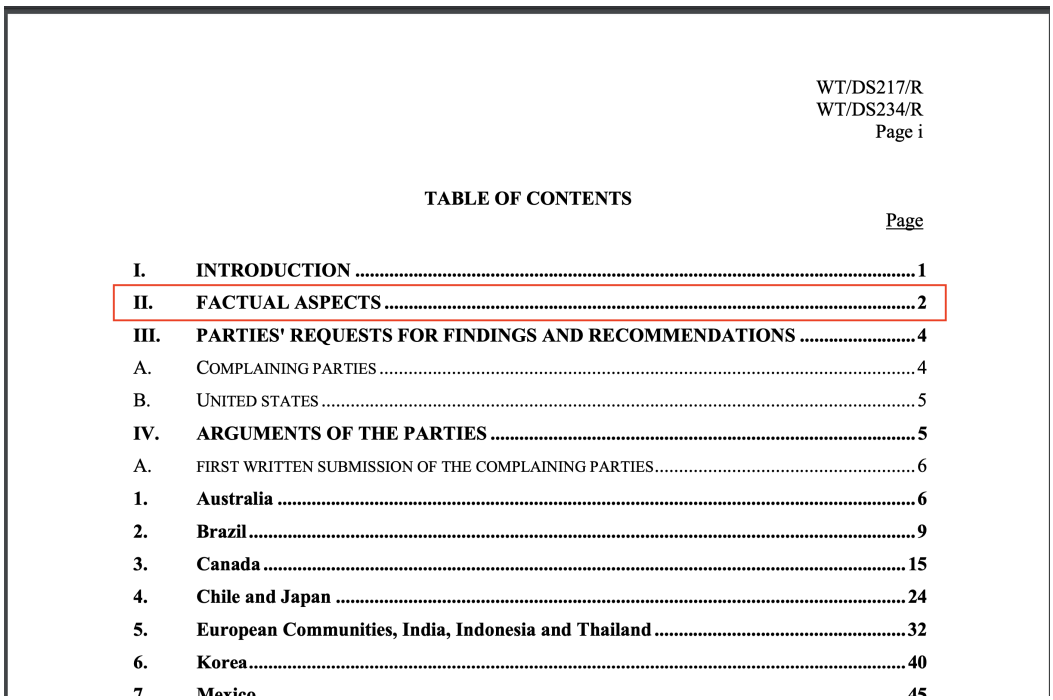
Table 1: Cited articles in *US - Offset (Byrd Amendment)* by complainants

Upon this understanding, I collected two different types of data for 143 different dispute cases requested to WTO DSB. (List of cases is available at Appendix A.2). One is textual descrip-

[1]Australia, Brazil, Chile, European Communities, India, Indonesia, Japan, Korea and Thailand

[2]It is worth noting that the WTO agreements comprises many different agreements covering each specific topic in trade such as *Agreement on Anti-dumping, Agreement on Subsidies and Countervailing Measures, Agreement on Agriculture* and so on.

tion of the dispute (*Check* the CDSOA example at Appendix A.1) and the other one is set of articles of the WTO agreements that are cited for each dispute (Appendix A.3). I will explain source, structure and collection method for two different types of data at the following subsections.



WT/DS217/R
WT/DS234/R
Page i

TABLE OF CONTENTS

Page

Figure 7: **Table of Contents of Panel Report:** Panel provides factual aspect in the panel report with its page location.

2.2 Factual Aspect: Textual Description of the Dispute

Textual description of the dispute is preferably called as *Factual Aspect* in WTO DSB. Since Panels always provide a factual aspect[3] that summarizes the content of the dispute in the panel report, I wrote a program that can automatically search and collect the panel reports from the WTO official doucment website[4]. Then I located the factual asepct using the page information inside the table of contents in the panel report as shown in Figure 7. By using this location, I excerpted factual aspect from 143 number of different panel reports listed in Figure 8.

[3]It's worth noting that Appellate Body doesn't provide any factual aspect because they use the factual aspect provided by the Panel.

[4]`http://docs.wto.org`

2.2.1 Joint Adjudication & Early Settlement

The number 143 seems small compared to the total number (596[5]) of cases requested to WTO DSB.

This is because, first, panel handles different cases together if the case is about the same trade policy. For example, in *US - Offset (Byrd Amendment)*, panel merged DS217[6] and DS234 together because they were asking the judicial opinion for the same government measure of the United States as shown in Figure 9. This paper selects the smallest case number as a representatitve number for this kind of joint Adjudication. For example, since DS217 and DS234 shares the same panel report, this paper choose DS217 as a representatitve number as shown in Figure 8 where the list includes DS217 but not DS234.

Secondly, members sometimes find *mutually agreeable solution* before the panel expresses its judicial opinion by publishing its panel report. Then Panel stops there and no factual aspect is available. This paper omitted this kind of early settled cases as well.

DS 2, 18, 22, 31, 34, 46, 56, 58, 60, 62, 67, 68, 69, 75, 76, 87, 90, 98, 103, 108, 121, 122, 135, 136, 139, 141, 146, 152, 155, 161, 162, 165, 166, 174, 175, 177, 184, 202, 207, 212, 217, 219, 221, 231, 234, 238, 244, 245, 246, 248, 257, 264, 265, 266, 267, 268, 269, 276, 282, 283, 286, 290, 294, 295, 296, 301, 302, 308, 312, 315, 316, 320, 321, 322, 332, 336, 339, 343, 344, 345, 350, 353, 360, 363, 366, 371, 379, 381, 384, 392, 394, 396, 397, 399, 400, 406, 412, 414, 415, 422, 425, 427, 429, 430, 431, 435, 436, 437, 440, 442, 447, 449, 453, 454, 456, 457, 461, 464, 468, 471, 472, 473, 475, 476, 477, 479, 480, 482, 483, 484, 485, 486, 488, 490, 492, 493, 495, 499, 504, 505, 513, 518, 523

Figure 8: **List of case number of collected panel reports:** "DS + number" uniquely identifies each dispute. For example, DS 523 refers to *US — Pipe and Tube Products (Turkey)* where United States was challenged by Turkey for its possibly inconsistent anti-dumping measure.

2.3 Cited Articles: Set of Articles Cited for the Same Dispute

Every lawsuit in WTO DSB cites multiple set of aritcles as shown in Table 1. To collect this set of articles claimed for the same dispute, I wrote a program that collects this set of articles cited for the same dispute from the WTO official webpage[7]. The webpage chronologically lists up all dispute cases requested to WTO DSB and the prgoram visits each page of 143 cases and collects

[5] As of November 1st, 2020.

[6] DS refers to Dispute Settelement. DS is official prefix that indicates the case in WTO DSB.

[7] `https://www.wto.org/english/tratop_e/dispu_e/dispu_status_e.htm`

WORLD TRADE ORGANIZATION

WT/DS217/R
WT/DS234/R
16 September 2002
(02-4742)

Original: English

UNITED STATES – CONTINUED DUMPING AND SUBSIDY OFFSET ACT OF 2000

Report of the Panel

Figure 9: **Cover of Panel Report:** Panel explicitly marks which different cases are handled together in the cover of the panel report. DS217 and DS234 are handled togehter in this example.

the cited articles. Among all the articles from different agreements of the WTO agreements[8], this paper collected articles from **General Agreement on Tariffs and Trade 1994 (GATT 1994)** only. This is because articles in GATT 1994 constitutes basic set of trade rules of WTO and other agreements elaborates the articles of GATT 1994 more in detail (World Trade Organization, 1999). For example, the official name of *Agreement on Anti-dumping* is *Agreement on Implementation of* ***Article VI of the GATT 1994*** where the name self-explains that it elaborates on the article VI of GATT 1994. The collected result is listed in the Appendix A.2. Figure 10 lists up 80 different articles of GATT 1994 cited in 143 cases without duplication.

2.3.1 Various Levels of Scope in Cited Articles

As shown in Figure 10, members sometimes cite articles in different levels of scope. For example, For the Article VI, member sometimes cites Article VI as a whole but sometimes cites Article VI:2 or Article VI:2(a). This is because two main judicial bodies of WTO DSB, *Panel and Appellate Body*, both constitute its legal precedents citing articles of the WTO agreements in various levels

[8]WTO agreeemnts is comprised of mutiple agreements such as General Agreement on Tariffs and Trade 1994, Agreement on Agriculture, Agreement on the Application of Sanitary and Phytosanitary Measures, Agreement on Textiles and Clothing, Agreement on Technical Barriers to Trade, Agreement on Trade-Related Investment Measures, Agreement on Implementation of Article VI of the General Agreement on Tariffs and Trade 1994 (antidumping), Agreement on Subsidies and Countervailing Measures, Agreement on Rules of Origin, Agreement on Safeguards and so on.

I, I:1, II, II:1, II:1(a), II:1(b), II:2, II:3, III, III:1, III:2, III:4, III:5, III:7, IV, IX, IX:2, V, V:1, V:2, V:3, V:3(a), V:4, V:5, V:6, V:7, VI, VI:1, VI:2, VI:2(a), VI:2(b), VI:3, VI:5(a), VI:6, VII, VII:1, VII:2, VII:5, VIII, VIII:1, VIII:3, VIII:4, X, X:1, X:2, X:3, X:3(a), XI, XI:1, XIII, XIII:1, XIII:2, XIII:3(b), XIX, XIX:1, XIX:2, XIX:3, XV, XVI, XVI:1, XVI:4, XVII, XVII:1, XVII:1(c), XVIII, XVIII:10, XVIII:11, XX, XXI, XXII, XXII:1, XXIII, XXIII:1, XXIII:1(a), XXIII:1(b), XXIV, XXIV:12, XXIV:5(b), XXIV:6, XXVIII

Figure 10: **Set of articles of GATT 1994 collected and used in this paper:** These articles comprises the node set V and their ordered pair comprises the edge set E in Figure 3

Let D is a set of DS case numbers listed in Figure 8.

Then there exists $c_d = \{v_d \in V \mid \ v_d \text{ is an article cited in the case } d\}\ \forall d \in D$

where V is set of articles listed in Figure 10.

Then define set of cited articles $C = \{c_d \mid d \in D\}$

Figure 11: **Formal Definition of Set of Cited Articles:** I formally define a set of cited articles C and the elements of C are listed in Appendix A.2.

of scope. Both judicial bodies cite the legal articles with the level of *Title, Article, Paragraph, Sentence* or *Term* as shown in Table 2. Following this jurisprudence, members also cite articles in different levels of scope to make their legal claim fit and valid according to the current jurisprudence of WTO DSB.

Table 2: **Various Levels of Scope Adopted to Cite Articles of WTO agreemnts**

Scope	**Quote**	**Source**
Title	"As the ***title*** **of Article 21 makes clear**, the task of panels . . . forms part of the process of the 'Surveillance of Implementation of the Recommendations and Rulings' of the DSB. . . ."	Appellate Body Report, *US – Shrimp (Malaysia)*, paras. 86-87.

Article	"The sequence of steps indicated above in the analysis of a claim of justification under **Article XX** reflects, not inadvertence or random choice, but rather the fundamental structure and logic of Article XX. . . ."	Appellate Body Report, *US – Shrimp (Malaysia)*, paras. 119-120.
Paragraph	"The verb 'may' in **Article VI:2** of the GATT 1994 is, in our opinion, properly understood as giving Members a choice between imposing an anti-dumping duty or not, as well as a choice between imposing an anti-dumping duty equal to the dumping margin or imposing a lower duty. . . ."	Appellate Body Report, *US – 1916 Act*, paras. 116.
Sentence	"The customary rules of interpretation of public international law as required by **the first sentence of Article 17.6(ii) of the Anti-Dumping Agreement**, do not admit of another interpretation as far as the issue of zeroing raised in this appeal is concerned."	Appellate Body Report, *US – Zeroing (EC)*, paras. 132-133.
Term	"Article II:1(a) provides that a Member shall accord to the 'commerce' of other Members treatment no less favourable than that provided for in its Schedule. **The term 'commerce'** is defined as referring broadly to the exchange of goods such that, in this provision, the 'commerce' of a Member should be understood to refer to all such exchanges of that Member"	Appellate Body Report, *Colombia – Textiles*, para. 5.34.

3 Methodology: Considerations and Development

This section introduces two main considerations to design the method used in this paper. Then it explains the method that is used to fit the network of articles of WTO agreement under those considerations.

3.1 Two Main Considerations For Design of Method

This paper considered two main points to determine its method to qualitatively fit the edge weights w^* for the *directed weighted graph* G^* as defined in Figure 3. One is importance of using the information represented in a form of textual description inside the content of dispute and legal article as exemplified in Appendix A.1 and Figure 6 respectively. The other one is about the way to generalize each member's startegic citation pattern. Since members of the WTO strategically cite the articles of WTO agreement expecting different outcomes thats serves member-specific national interest (Johns and Pelc, 2014; Pelc, 2014; Strezhnev, 2014), this paper selected a method that can generalize this member speicific citation pattern. These two considerations and the solution will be explained in the following subsections.

3.1.1 Importance of Using Textual Information

This paper emphasizes the necessity of using textual information to qualitatively fit the edge weights w^* for the *directed weighted graph* G^* that is defined in Figure 3. One can simply consider a co-citation pattern between the articles of WTO agreements as a a regulatory system of WTO DSB, however, it simply allocates a huge edge weight for frequently cited articles and fails to explain how articles interact to achieve main principles of WTO as exemplified in Figure 2.

This failure is mainly due to the insufficient information in co-citation matrix. Members tend to cite the articles of the WTO agreements based on the complex characteristics of the trade policy that led to the dispute, however, co-citation pattern omits this prior information.

To emphasize the importance of using textual information, I prepared two differeent matrices $W_{\text{co-cites}}$ and W_{text} that is following the definition of *edge weight matrix* W in Figure 3. $W_{\text{co-cites}}$ is calculated only using the co-citation pattern between articles of the WTO agreements as formally defined in Figure 14. W_{text} is the one that is fitted using the textual information and the way how it's fitted will be introduced at the following body in this section. Two Heatmaps visualized in Figure 12 shows how sparse the $W_{\text{co-cites}}$ is. This sparisty led to the insufficient information to qualitatively maps the regulatory system of WTO DSB. In contrast with it, if we fit the *edge weight*

matrix W using the textual information, we get more dense matrix as visualized in Figure12(b).

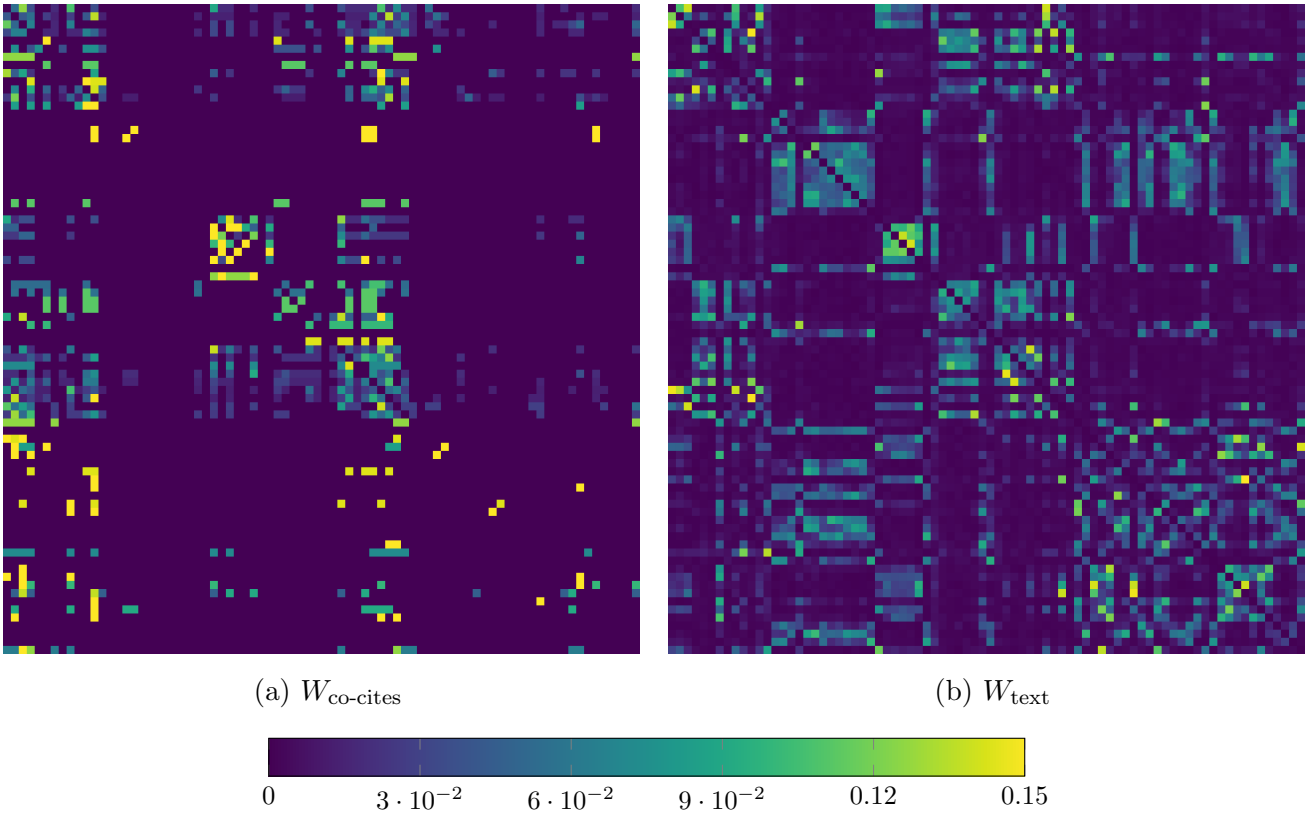


Figure 12: **Heatmap of Two Different Edge Weight Matrices:** Above two subfigures visualizes two different *edge weight matrices* $W_{\text{co-cites}}$ and W_{text}. One can check that $W_{\text{co-cites}}$ is sparser than W_{text}.

Upon this observation, this paper justifies the use a deep neural network to process information embedded in the text description of the trade policy and legal articles. This is because deep neural network is known to extract information from the text effectively to perform various tasks such as text classification (Minaee et al., 2020), text summarization (Magdum and Rathi, 2021) and text generation (Guo et al., 2017) and so on.

3.1.2 Generalization of Each Member's Strategic Citation

This paper aims to map the regulatory system of WTO DSB in a form of *directed weighted graph* G as defined in Figure 3. To achieve this purpose, we need to fit G^* to generalize member specific strategic citation behavior. In terms of generalization, deep neural network is known to generalize well despite its large capacity (Neyshabur et al., 2017), possible instability of training algorithm (Charles and Papailiopoulos, 2017), nonrobustness (Zahavy et al., 2017), and sharp minima (Dinh et al., 2017). Therefore this paper trains a deep neural network without any member specific information such as geolocation, GDP or specialized industry. By training a deep

Let δ_{ij}^d is defined to be 1 if $\{(v_i, v_j) \mid v_i, v_j \in V \text{ and } i \neq j\} \subset c_{d \in D}$ else 0

where V, D and c_d is defined as in Figure 11.

Then let *co-citation matrix* $M = (m_{ij}) \in \mathbb{N}^{|V| \times |V|}$ s.t. $m_{ij} = \sum_{d \in D} \sum_{i,j \in V} \delta_{ij}^d$

(a) **Formal Definition of Co-citation Matrix**

	I	**I:1**	**II**	**II:1**	$\cdots$
I	**0**	**3**	**7**	**2**	
I:1	**3**	**0**	**3**	**4**	
II	**7**	**3**	**0**	**4**	
II:1	**2**	**4**	**4**	**0**	
$\vdots$					

(b) **Illustration of Co-citation Matrix**

Figure 13: **Formal Definition and Illustration of Co-citation Matrix:** This paper defines co-citation matrix M as subfigure (a) and it's illustrated as subfigure (b) using the paper's dataset. Note that co-citation matrix is *symmetric*, $m_{ij} = m_{ji}$ $\forall i, j \in V$.

neural network only using a description of the possible inconsistent trade policy and the legal text of the cited articles, this paper expects a fitted G^* can show interactions between articles of the WTO agreements without being biased to member specific strategic citation.

3.2 Design of Deep Neural Network

Upon the justification of using deep neural network with two main reasons, I introduce this paper's design of deep neural network that is expected to efficiently encode members' citation pattern for given possible inconsistent trade policy.

3.2.1 Visualization of How Member Cites in WTO DSB

For given M defined in Figure 13(a),

let *normalized co-citation matrix* $N = (n_{ij}) \in \mathbb{R}^{|V|\times|V|}$ s.t. $n_{ij} = \dfrac{m_{ij}}{\sum_{j\in V} m_{ij}}$

(a) **Formal Definition of Normalized Co-citation Matrix**

	I	**I:1**	**II**	**II:1**	$\cdots$	
I	**0**	**0.053**	**0.125**	**0.035**		$\sum_{j\in V} n_{ij} = 1$
I:1	**0.040**	**0**	**0.04**	**0.054**		
II	**0.114**	**0.049**	**0**	**0.065**		
II:1	**0.032**	**0.065**	**0.065**	**0**		
$\vdots$						

(b) **Illustration of Noramlized Co-citation Matrix**

Figure 14: **Formal Definition and Illustration of Normalized Co-citation Matrix:** This paper defines normalized co-citation matrix N of M as subfigure (a) and it's illustrated as subfigure (b) using the paper's dataset. Note that normalized co-citation matrix is no more *symmetric*, $n_{ij} \neq n_{ji}\ \forall i, j \in V$. This definition is prepared to fit the definition of co-citation matrix to that of W in Figure 3.

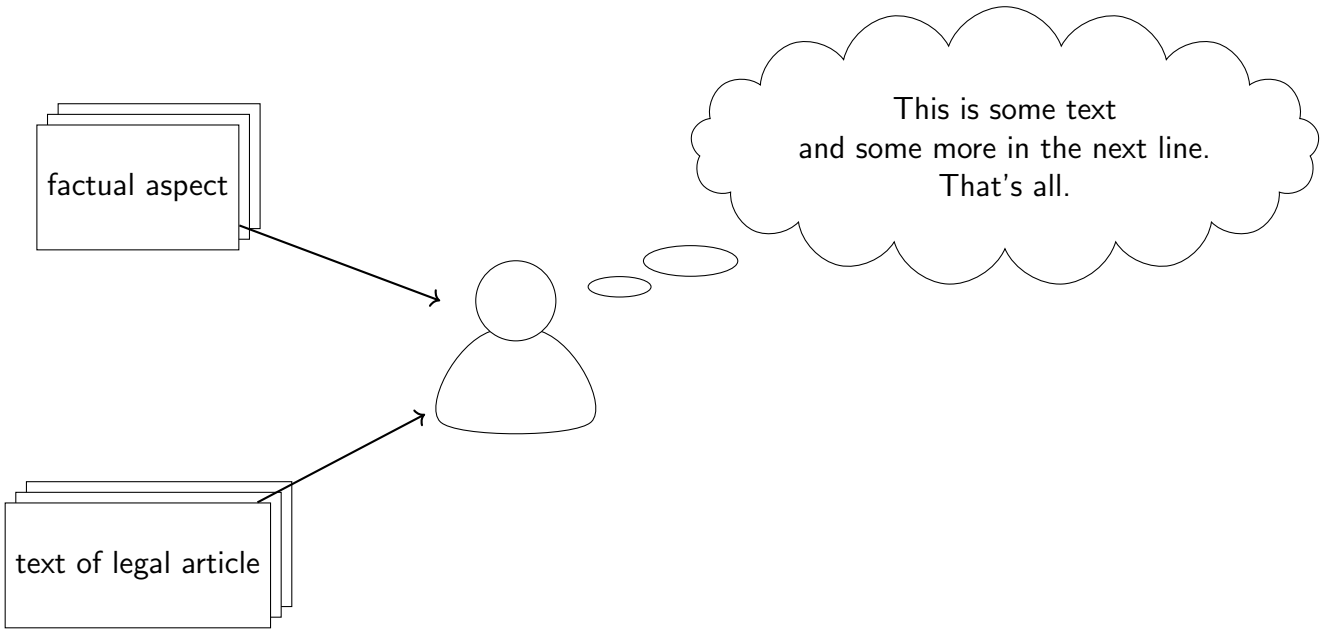


Figure 15: **Visualization of How Member Cites in WTO DSB:** With two different contexts, factual aspect and text of legal articles, a member can judge whether the given legal article is appropriate to be cited or not.

References

Charles, Zachary and Dimitris Papailiopoulos. 2017. "Stability and Generalization of Learning Algorithms that Converge to Global Optima.".

Dinh, Laurent, Razvan Pascanu, Samy Bengio and Yoshua Bengio. 2017. "Sharp Minima Can Generalize For Deep Nets.".

Florini, Ann. 1996. "The Evolution of International Norms." *International Studies Quarterly* 40(3):363–389.

Fredebeul-Krein, Markus and Andreas Freytag. 1999. "The case for a more binding WTO agreement on regulatory principles in telecommunication markets." *Telecommunications Policy* 23(9):625 – 644.

Guo, Jiaxian, Sidi Lu, Han Cai, Weinan Zhang, Yong Yu and Jun Wang. 2017. "Long Text Generation via Adversarial Training with Leaked Information.".

Hufbauer, Gary and Sherry Stephenson. 2007. "Services Trade: Past Liberalization and Future Challenges." *Journal of International Economic Law* 10(3):605–630.

Huynh-Thu, Vân Anh, Alexandre Irrthum, Louis Wehenkel and Pierre Geurts. 2010. "Inferring Regulatory Networks from Expression Data Using Tree-Based Methods." *PLOS ONE* 5(9):1–10.

Johns, Leslie and Krzysztof J. Pelc. 2014. "Who Gets to Be In the Room? Manipulating Participation in WTO Disputes." *International Organization* 68(3):663–699.

Magdum, P. G. and Sheetal Rathi. 2021. A Survey on Deep Learning-Based Automatic Text Summarization Models. In *Advances in Artificial Intelligence and Data Engineering*, ed. Niranjan N. Chiplunkar and Takanori Fukao. Singapore: Springer Singapore pp. 377–392.

Minaee, Shervin, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu and Jianfeng Gao. 2020. "Deep Learning Based Text Classification: A Comprehensive Review.".

Neyshabur, Behnam, Srinadh Bhojanapalli, David McAllester and Nathan Srebro. 2017. "Exploring Generalization in Deep Learning.".

Palmeter, D. and P.C. Mavroidis. 2004. *Dispute Settlement in the World Trade Organization: Practice and Procedure*. Dispute Settlement in the World Trade Organization: Practice and Procedure Cambridge University Press.

Pelc, Krzysztof J. 2014. "The Politics of Precedent in International Law: A Social Network Application." *The American Political Science Review* 108(3):547–564.

Shaffer, Gregory. 2004. *Power, governance, and the WTO: a comparative institutional approach.* Cambridge Studies in International Relations Cambridge University Press p. 130–160.

Strezhnev, BuenoAnton. 2014. "Using Latent Space Models to Study International Legal Precedent: An Application to the WTO Dispute Settlement Body." *American Political Science Association 2014 Annual Meeting* .

World Trade Organization. 1999. *WTO Agreements Series.* Number no. 2 *in* "WTO Agreements Series" World Trade Organization.

World Trade Organization. 2017. *A Handbook on the WTO Dispute Settlement System.* A WTO Secretariat publication Cambridge University Press.

Zahavy, Tom, Bingyi Kang, Alex Sivak, Jiashi Feng, Huan Xu and Shie Mannor. 2017. "Ensemble Robustness and Generalization of Stochastic Deep Learning Algorithms.".