Pianzola, F. (2021). Modeling Literary Criticism: How to Do It and How to Teach It to Humans and Machines. *Style*, *55*(1), 111–117. <https://doi.org/10.5325/style.55.1.0111>

**Preprint version**

How would you replicate Eric Auerbach’s masterful critical readings of Western literature? How to teach students effectively and in detail what steps to follow in analyzing a literary text? Is it possible to teach a computer to perform a thematic critical reading of a novel or short story? Dan Shen does not address these questions explicitly, but the methodology she uses to develop a theory of dual narrative dynamics invites to explore them. In the book *Style and Rhetoric of Short Narrative Fiction. Covert Progressions Behind Overt Plots*, Shen already extensively illustrated her analytical method, taking into account concepts from narratology and stylistics that she put at work in synergy. The essay “‘Covert Progression’ and Dual Narrative Dynamics” is a step further in clarifying how decades of literary criticism can be modelled into an “interpretation protocol” (Liu) useful to start exploring the complexity of literary texts. A model that could inform the learning process of both humans (cf. Herman et al.) and machines (cf. McCarty; Flanders and Jannidis).

Shen explicitly presents interpretive processes that go beyond textual elements, linking them to historical context, authorial traits, and literary conventions. Accounting for these elements in close reading is often a challenge for both students and computers alike, because they require a knowledge extending beyond the boundaries of the text under scrutiny. In other words, having access to the right kind of contextual information and making meaningful connections. For students, this means spending a lot of time reading secondary sources and creating mental maps that will help them interpret the text. For computers, it means being fed additional corpora of texts and a set of rules to process them when required, based on certain features of the analyzed text or questions posed by the critic. In both cases, a good literary theory is essential in order to know what kind of information to look for and how to use it.

Similarly to how folklore studies and narratology made explicit and formalized the dynamics of plot, “covert progression” is a useful model for thematic interpretation. I think this is a very positive aspect of Shen’s theory and I would like to summarize her model in more pedagogical terms – departing from her fifteen theses – in order to show how it can be useful for teaching step-by-step how to perform a thematic critical reading of a literary text. Interestingly, Shen’s model can be used not only with students but with computers as well, providing an operative model translatable into instructions for machine reading (Hayles; Kestemont and Herman).

 The steps to be followed are:

1. individuate odd, puzzling, trivial or digressive linguistic choices (“recalcitrant materials,” Phelan), or peripheral details;
2. check if they appear throughout the text (this process may require many rereadings because newfound material can suggest looking at other aspects);
3. if yes, outline the event structure of the covert progression, following the appearance of the recalcitrant materials;
4. consider why the identified elements are not part of the overt plot;
5. find contextual explanations (authorial, historical, intertextual) that can guide the interpretation.

When finding recalcitrant materials,

one need examine carefully whether they interact with other linguistic choices in different parts of the text to convey implicitly a contrastive or even opposite kind of thematic meaning and to portray different character images. If the result of the exploration is more or less validating, the analyst not only needs to carry out the stylistic analysis along the two different trajectories of signification (frequently adjusting the direction in the process of the analysis), but also needs to pay attention to the interaction between them. (Shen, “‘Covert Progression’ and Dual Narrative Dynamics”)

 Once the probability of having dual narrative dynamics is assessed, the analyst has to explore how they interact with each other – tending either towards harmonious complementation or drastic subversion – and remembering that “‘covert progression’ is an undercurrent *paralleling* the plot development, with the *same* central character(s) and the *same* events but implicitly conveying contrastive or opposite thematic significance” (Shen, “‘Covert Progression’ and Dual Narrative Dynamics”). In order to do so, Seymour Chatman’s distinction between “plot of resolution” and “plot of revelation” can be useful: cases in which events are causally related and progress towards a denouement, marked by a process of change, vs. cases in which events are not resolved (happily or tragically), but rather a state of affairs is revealed (Chatman 45–48). This distinction can guide questions that will help to outline dual dynamics with respect to various narrative features and create specific models for each of them.

A model of dual event structure:

(i) *Event structure of the overt plot:*What constitutes the conflict and tension of the events in the plot? Are the events revelatory or moving towards a resolution?

(ii) *Event structure of the covert progression:* What constitutes the conflict and tension of the events in the covert progression? Are the events revelatory or moving towards a resolution?

 (iii) *Relationship between the two:* Are they subversive or complementary to each other?

A model of dual characterization and character image:

(i) *Characterization and character image in the plot development:* What features of a character are (emphatically) depicted in the overt plot? What image of the character emerges in this narrative movement?

(ii) *Characterization and character image in the covert progression:* What traits of the character are (emphatically) conveyed in the covert progression? What image of the character comes out of this undercurrent?

 (iii) *Relationship between the two:* Are they subversive or complementary to each other?

A model of dual unreliability:

(i) *Reliability of the narrator in the plot development*: How reliable is the narrator in the overt plot? What is the criterion for judging the narrator’s unreliability in this narrative movement?

(ii) *Reliability of the narrator in the covert progression*:How reliable is the narrator in the covert progression? What is the criterion for judging the narrator’s unreliability in this narrative movement?

 (iii) *Relationship between the two*: Are they subversive or complementary to each other?

As suggested by Shen, along the same line, models for other narrative instances can be developed: dual implied author, dual authorial audience, dual narrative distance, dual focalization, dual narrative tone, dual model of story and discourse.

 To differentiate between plots of resolution and plots of revelation various methods can be tentatively combined: character identification (Bamman et al.), events detection (Sims et al.; Gius et al.), motifs extraction (Yarlott and Finlayson), evaluation of narrative tellability (Berov), and analysis of how the frequency of words varies throughout the story, based on their syntactic, semantic, or narrative function (McClure; Boyd et al.), but also on their emotional valence (Archer and Jockers; Reagan et al.).

Then, it would not be too difficult to code steps 1 to 3. Recalcitrant materials can be identified with word embeddings (Xu et al.) and topic modelling techniques (Jockers; Algee-Hewitt, Heuser, and Moretti), that is determining the semantic or topic similarity of words based on the probability with which they co-occur in the text. Alternatively, and more transparently, we could explore terms frequencies in a text with respect to their frequency in a reference corpus (TF-IDF), words frequently recurring in the textual vicinity of a target term (collocations), and groups of terms often occurring together (n-grams). Learning how to focus the analysis on different narrative instances, like implied author or focalization, is a more complex task for a computer, but there is some progress in this direction (Reiter et al.; Kim and Klinger; Finlayson; Min and Park). Interestingly, preliminary manual annotation of texts by experts or lay readers is required in this kind of supervised machine learning approaches, so that the computer can identify as closely as possible what the literary critic wants. Thus, Shen’s model can also help in teaching humans how to teach machines. Steps 4 and 5, the contextual explanation, can only be partly automated, by including a reference corpus made of other texts by the same author, genre, or historical context. However, ultimately it is always up to human readers to evaluate the more plausible interpretation.

More than any hyper-technological application of machine reading, new advances in computer-assisted critical reading need good theories that can guide and orient automatization, as well as interpretation protocols that makes the hermeneutic process replicable and more easily debatable (WE1S). And new advances in critical reading and literary theory would benefit from research apprentices that can work with well-designed procedures and focus their analytical and theoretical skills on interpretative processes, not on the collection of data scattered throughout the text (or hundreds of texts). Dan Shen’s model of dual narrative dynamics is an excellent candidate for both purposes and I look forward to seeing it integrated into courses teaching narrative theory and into computational analyses of literary texts.

**References**

Algee-Hewitt, M., Heuser, R. and Moretti, F. ‘On Paragraphs. Scale, Themes, and Narrative Form’. *Pamphlets of the Standford Literary Lab*, vol. 10, 2015.

Archer, Jodie, and Matthew L. Jockers. *The Bestseller Code: Anatomy of the Blockbuster Novel*. St. Martin’s Press, 2016.

Bamman, David, et al. ‘A Bayesian Mixed Effects Model of Literary Character’. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics, 2014, pp. 370–79. doi:10.3115/v1/P14-1035.

Berov, Leonid. ‘Towards a Computational Measure of Plot Tellability’. *AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2017. https://aaai.org/ocs/index.php/AIIDE/AIIDE17/paper/view/15886.

Boyd, Ryan L., et al. ‘The Narrative Arc: Revealing Core Narrative Structures through Text Analysis’. *Science Advances*, vol. 6, no. 32, Aug. 2020, eaba2196. doi:10.1126/sciadv.aba2196.

Chatman, Seymour. *Story and Discourse. Narrative Structure in Fiction and Film.* Cornell University Press, 1978.

Finlayson. ‘Inferring Propp’s Functionsfrom Semantically Annotated Text’. *The Journal of American Folklore*, vol. 129, no. 511, 2016, pp. 55–77. doi:10.5406/jamerfolk.129.511.0055.

Flanders, Julia, and Fotis Jannidis, editors. *The Shape of Data in Digital Humanities: Modeling Texts and Text-Based Resources*. Routledge, 2018.

Gius, Evelyn, et al. ‘Detection of Scenes in Fiction’. *DH2019*, 2019, https://dev.clariah.nl/files/dh2019/boa/0608.html.

Hayles, Katherine. *How We Think: Digital Media and Contemporary Technogenesis*. The University of Chicago Press, 2012.

Herman, David, et al., editors. *Teaching Narrative Theory*. Modern Language Association of America, 2010.

Jockers, Matthew L. *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press, 2013.

Kestemont, Mike, and Luc Herman. ‘Can Machines Read (Literature)?’ *Umanistica Digitale*, no. 5, 2019. doi:10.6092/issn.2532-8816/8511.

Kim, Evgeny, and Roman Klinger. ‘Who Feels What and Why? Annotation of a Literature Corpus with Semantic Roles of Emotions’. *Proceedings of the 27th International Conference on Computational Linguistics*, Association for Computational Linguistics, 2018, pp. 1345–539. https://www.aclweb.org/anthology/C18-1114.

Liu, Alan. *Humans in the Loop: Humanities Hermeneutics and Machine Learning*. Zentrum für Informations- und Medientechnologien (IMT) der Universität Paderborn, 2020. https://www.youtube.com/watch?v=lnfeOUBCi3s&feature=youtu.be.

McCarty, Willard. ‘Modeling: A Study in Words and Meanings’. *A Companion to Digital Humanities*, edited by Susan Schreibman et al., Blackwell, 2004.

McClure, David. ‘A Hierarchical Cluster of Words across Narrative Time – Stanford Literary Lab’. *Stanford Literary Lab*, 31 July 2017, https://litlab.stanford.edu/hierarchical-cluster-across-narrative-time/.

Min, Semi, and Juyong Park. ‘Modeling Narrative Structure and Dynamics with Networks, Sentiment Analysis, and Topic Modeling’. *PLOS ONE*, vol. 14, no. 12, Dec. 2019, e0226025. doi:10.1371/journal.pone.0226025.

Reagan, Andrew J., et al. ‘The Emotional Arcs of Stories Are Dominated by Six Basic Shapes’. *EPJ Data Sci.*, vol. 5, 2016, pp. 5–31, doi:10.1140/epjds/s13688-016-0093-1.

Reiter, Nils, et al. ‘A Shared Task for the Digital Humanities Chapter 1: Introduction to Annotation, Narrative Levels and Shared Tasks’. *Journal of Cultural Analytics*, 2019. doi:10.22148/16.048.

Shen, Dan. *Style and Rhetoric of Short Narrative Fiction. Covert Progressions behind Overt Plots*. Routledge, 2014.

Sims, Matthew, et al. ‘Literary Event Detection’. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, 2019, pp. 3623–34. doi:10.18653/v1/P19-1353.

WE1S. ‘Topic Model Interpretation Protocol’. *WE1S*, 26 June 2019. *we1s.ucsb.edu*, https://we1s.ucsb.edu/research/we1s-methods/topic-model-interpretation-protocol/.

Xu, Huimin, et al. ‘The Cinderella Complex: Word Embeddings Reveal Gender Stereotypes in Movies and Books’. *PLOS ONE*, vol. 14, no. 11, Nov. 2019, e0225385. doi:10.1371/journal.pone.0225385.

Yarlott, W. Victor H., and Mark A. Finlayson. ‘Learning a Better Motif Index: Toward Automated Motif Extraction’. *7th Workshop on Computational Models of Narrative (CMN 2016)*, edited by Ben Miller et al., Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik GmbH, Wadern/Saarbruecken, Germany, 2016, pp. 7:1-7:10, doi:10.4230/OASIcs.CMN.2016.7.