



THE STEIN FACULTY OF COMPUTER  
AND INFORMATION SCIENCE  
Ben-Gurion University of the Negev

# Structural Patterns in Award-Winning Novellas

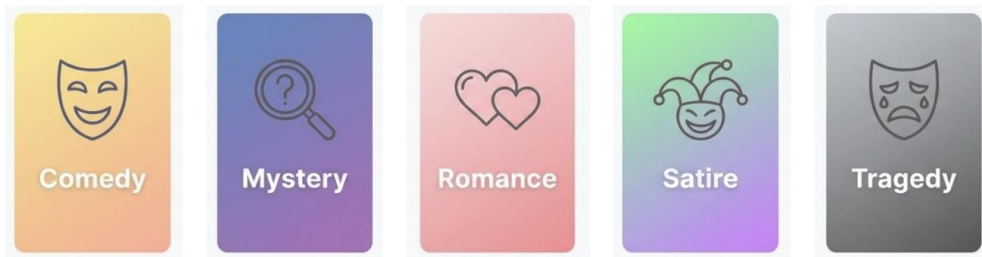
---

Milka Kaplan, Armin Shmilovici, Mark Last






# Motivation

PatternTeller ([Lima et al. 2025](#)) describes five genre-specific narrative patterns:



## Research Questions

-  Do literary models (e.g. PatternTeller) manifest in award-winning novellas ?
-  Which structural elements in a literary model are most genre-distinctive in novellas?
-  Can LLMs reliably annotate these patterns?

Example of the pattern Tragedy:

**STAGE 1**  
**Stage of normalcy**  
The protagonist lives an ordinary life, unaware of the challenges they will soon face.

**STAGE 2**  
**Catalyst**  
A disruptive event or piece of information catalyzes a change in the protagonist's status quo.

**STAGE 3**  
**Awakening**  
The protagonist begins to question their life and values — a shift in perspective begins the quest.

**STAGE 4**  
**Confrontation**  
Internal or external conflicts challenge the protagonist's ambitions and moral compass.

**STAGE 5 — TURNING POINT**  
**Crisis**  
A pivotal decision that could alter the entire course of the protagonist's journey. No return.

**STAGE 6**  
**Downfall**  
A deep loss or realization — the direct consequence of earlier decisions — leads to the fall from grace.

**STAGE 7**  
**Final act**  
The protagonist faces the ultimate consequences: demise or a profound, irreversible change.

# PatternTeller's Narrative Structures

## Comedy



- 1 Set Up
- 2 Catalyst
- 3 Rising Action
- 4 Climax
- 5 Falling Action
- 6 Resolution
- 7 Denouement



7 stages

## Mystery



- 1 Intro & Setting
- 2 Mystery Presented
- 3 Investigation
- 4 Uncovering Clues
- 5 Confronting Suspects
- 6 Red Herrings
- 7 Realization
- 8 Final Confrontation
- 9 Resolution



9 stages

## Romance



- 1 Call to Adventure
- 2 Meeting Mentor
- 3 Departure
- 4 Trials
- 5 Inmost Cave
- 6 Ordeal
- 7 Reward
- 8 Return
- 9 Resurrection
- 10 Return with Elixir



10 stages

## Satire



- 1 Setting & Characters
- 2 Revelation of System
- 3 Awakening
- 4 Resistance
- 5 Crisis & Struggle
- 6 Consequence
- 7 Final Revelation
- 8 Closure



8 stages

## Tragedy










- 1 Normalcy
- 2 Catalyst
- 3 Awakening
- 4 Confrontation
- 5 Crisis
- 6 Downfall
- 7 Final Act



7 stages

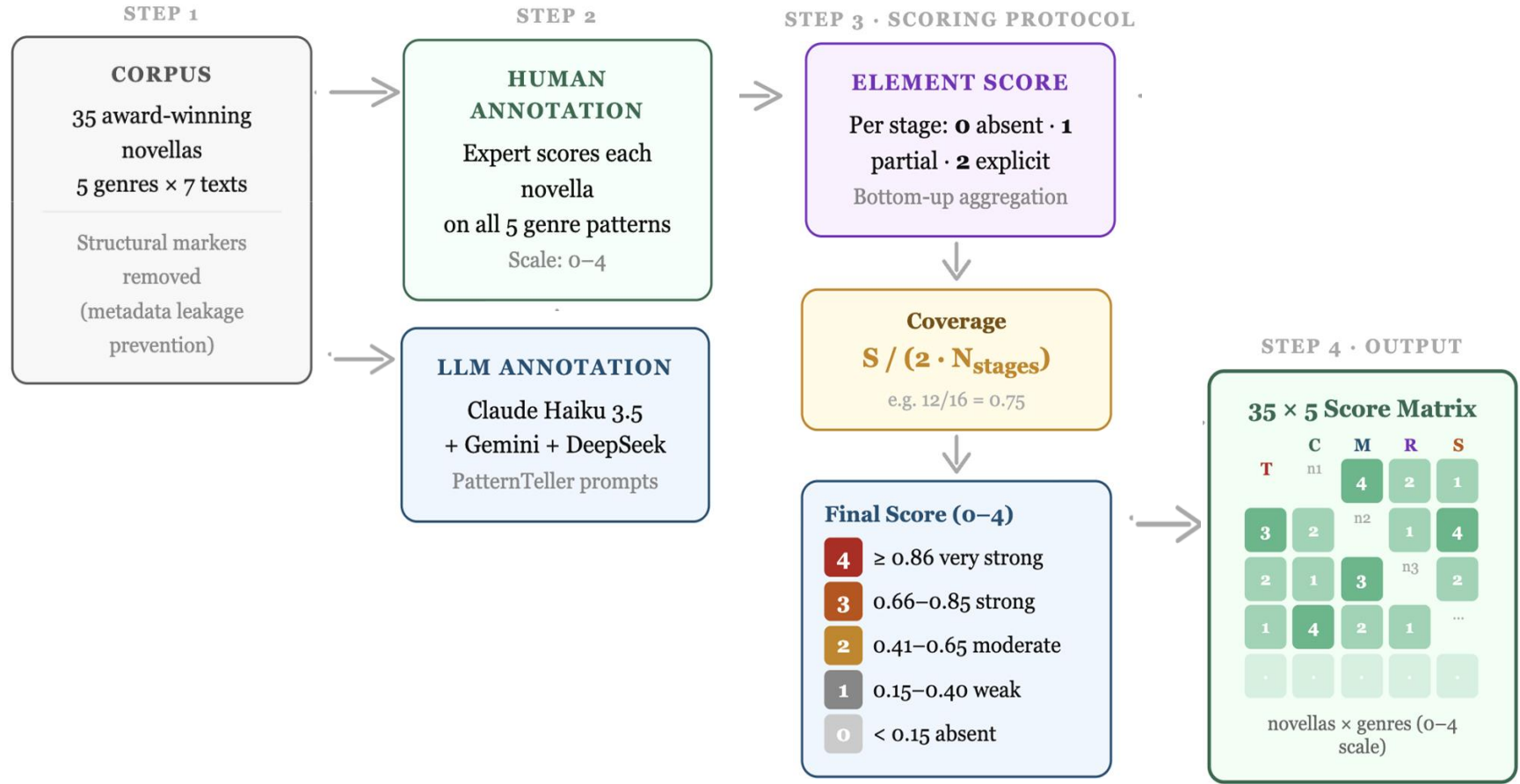
# Dataset Description

35 highly rated English novellas selected from the [Goodreads "World's Greatest Novellas"](#) list. (ranging 7,500-102,810 words)

goodreads		Search books
		All Votes Add Books To This List
1	 <b>Animal Farm</b> by George Orwell ★★★★★ 4.02 avg rating — 4,621,496 ratings score: 51,448, and 522 people voted	Want to Read Rate this book ★★★★★
2	 <b>Of Mice and Men</b> by John Steinbeck ★★★★☆ 3.90 avg rating — 2,866,249 ratings score: 31,319, and 323 people voted	Want to Read Rate this book ★★★★★
3	 <b>The Metamorphosis</b> by Franz Kafka ★★★★☆ 3.91 avg rating — 1,459,300 ratings score: 27,643, and 284 people voted	Want to Read Rate this book ★★★★★
4	 <b>The Old Man and the Sea</b> by Ernest Hemingway ★★★★☆ 3.81 avg rating — 1,328,777 ratings score: 26,166, and 268 people voted	Want to Read Rate this book ★★★★★
5	 <b>The Stranger</b> by Albert Camus ★★★★★ 4.03 avg rating — 1,432,629 ratings score: 23,441, and 241 people voted	Want to Read Rate this book ★★★★★
6	 <b>The Little Prince</b> by Antoine de Saint-Exupéry ★★★★★ 4.33 avg rating — 2,513,000 ratings score: 20,284, and 210 people voted	Want to Read Rate this book ★★★★★
7	 <b>A Christmas Carol</b> by Charles Dickens ★★★★★ 4.09 avg rating — 946,876 ratings score: 17,574, and 182 people voted	Want to Read Rate this book ★★★★★

Genre	Corpus filename	Year	Author	Word count	Avg. rating	Goodreads score
tragedy	002_mice-and-men.txt	1937	John Steinbeck	29673	3.89	31023
	302_the_great_gatsby.txt	1925	F. Scott Fitzgerald	48447	3.93	168
	000_the_old_man_and_the_sea.txt	1952	Ernest Hemingway	26269	3.81	26163
	001_animal-farm.txt	1945	George Orwell	29908	3.99	51254
	051_the_picture_of_dorian_gray.txt	1891	Oscar Wilde	54348	4.13	2753
	028_Lord_of_the_flies.txt	1954	William Golding	59725	3.70	4283
	093_to_kill_a_mocking_bird.txt	1960	Harper Lee	99235	4.26	1034
satire	067_the_catcher_in_the_rye.txt	1951	J.D. Salinger	74190	3.80	1684
	102_Miss_Lonelyhearts.txt	1933	Nathanael West	23203	3.75	941
	109_being_there.txt	1970	Jerzy Kosinski	25149	3.86	842
	156_the_13_clocks.txt	1950	James Thurber	11750	4.00	457
	258_farmer_giles_of_ham.txt	1949	J.R.R. Tolkien	16840	3.88	209
	317_the_wonderful_o.txt	1957	James Thurber	10757	3.87	152
	407_the_restaurant_at_the_end.txt	1980	Douglas Adams	52256	4.20	100
romance	055_giovannis_room.txt	1956	James Baldwin	54693	4.33	2647
	151_on_chesil_beach.txt	2007	Ian McEwan	36743	3.61	475
	270_the_governess_affair.txt	2012	Courtney Milan	32960	3.85	195
	377_greater_than_rubies.txt	2012	Hallee Bridgeman	56920	4.04	100
	418_breathless.txt	2012	Sophie Jordan	25116	3.80	99
	489_skin_deep.txt	2012	Pamela Clare	52170	4.12	94
	560_how_to_marry_a_werewolf.txt	2018	Gail Carriger	47391	4.19	81
mystery	054_doyle-hound-383.txt	1902	Arthur Conan Doyle	59321	4.13	2678
	135_the_maltese_falcon.txt	1930	Dashiell Hammett	66479	3.87	553
	036_the_turn_of_the_screw.txt	1898	Henry James	39471	3.39	3488
	191_every_heart_a_doorway.txt	2016	Seanan McGuire	38579	3.72	367
	337_the_library_window.txt	1896	Mrs. Oliphant	18403	3.71	137
	345_lyras_oxford.txt	2003	Philip Pullman	7354	3.62	132
	495_hammers_on_bone.txt	2016	Cassandra Khaw	17662	3.62	93
comedy	043_the_prime_of_miss_jean_brodie.txt	1961	Muriel Spark	39498	3.70	3121
	079_shopgirl.txt	2000	Steve Martin	36583	3.42	1316
	316_the_canterville_ghost.txt	1887	Oscar Wilde	11373	3.88	154
	046_cannery_row.txt	1945	John Steinbeck	49837	4.06	2958
	268_under_the_greenwood_tree.txt	1872	Thomas Hardy	56707	3.65	196
	452_the_perilous_life_of_jade.txt	2012	Zen Cho	24874	3.79	97
	201_the_tiny_wife.txt	2010	Andrew Kaufman	14313	3.61	343

# Methodology for Evaluating Literary Models



# Sample LLM Analysis Output

FILE: Copy of Copy of 1\_animal-farm.txt

PROCESSED: 2026-01-14 12:03:26

WORDS: 30,050

=====  
I'll analyze this text using the Tragedy pattern.

=== MODEL EVALUATIONS ===

Model: Tragedy

StageTotal: 12/14

Coverage: 0.86

Score: 4

Justification: The text perfectly follows the tragic arc of a revolution betrayed, where the protagonist (the animal collective) experiences a complete moral and social downfall from their initial idealistic hopes.

Stage Breakdown:

1. Normalcy (2/2): Animals live under Jones's oppressive farm regime
2. Catalyst (2/2): Old Major's revolutionary speech inspires rebellion
3. Awakening (2/2): Animals overthrow Jones and establish Animal Farm
4. Confrontation (2/2): Internal conflicts between Snowball and Napoleon emerge
5. Crisis (2/2): Napoleon consolidates power through manipulation and violence
6. Downfall (1/2): Animals gradually lose their revolutionary principles
7. Final Act (1/2): Pigs become indistinguishable from humans, completing the tragic transformation

The text demonstrates a comprehensive tragic structure where the initial noble ideals are systematically corrupted, leading to a complete betrayal of the original revolutionary vision. The animals' dream of equality is ultimately destroyed by the very leaders who promised liberation.

The raw output produced by Claude the analysis pipeline for Animal Farm (Orwell, 1945).

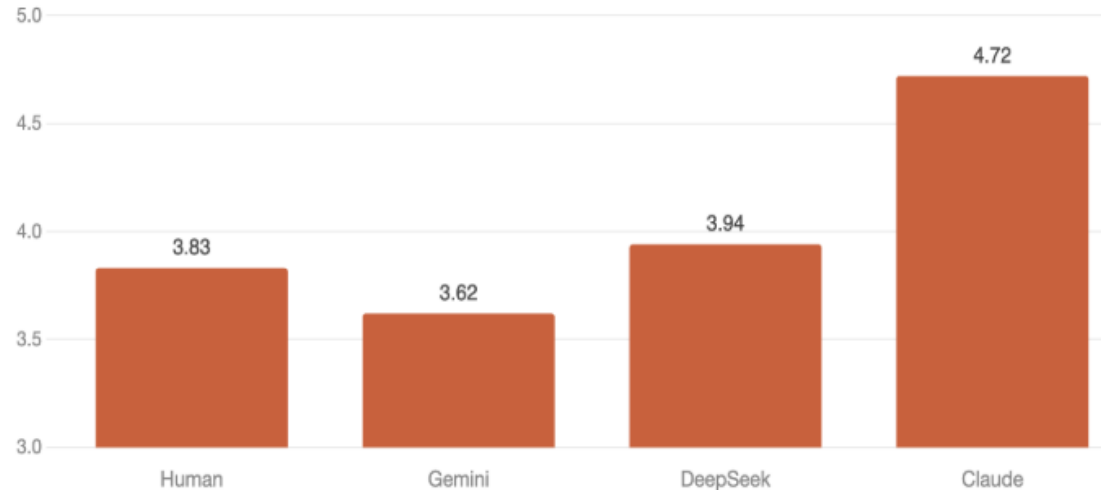
# Herfindahl–Hirschman Index (HHI) for different LLMs

HHI - to determine the concentration of genres within a single novella

$$HHI = \sum_{i=1}^N \left( \frac{s_i}{\sum_{j=1}^N s_j} \right)^2$$

where  $s_i$  is the score assigned to an element of the genre  $i$ , and  $N = 5$  is the number of genres.

Effective Number of Genres ( $N_{eff} = 1/HHI$ )



*Low value is better and demonstrate an effective number of genres*

**Among all LLMs, Gemini is the best based on the Neff of HHI.**

# Selection of Key Narrative Elements

## The most genre-distinctive structural elements

### Comedy

Stage	Concept	Rate	Imp.
Stage 4	<b>Climax</b>	1.00	0.918
Stage 5	<b>Falling Action</b>	1.00	0.884
Stage 6	<b>Resolution</b>	1.00	0.782
Stage 7	<b>Denouement</b>	1.00	0.781

### Mystery

Stage	Concept	Rate	Imp.
Stage 2	<b>Presentation of Mystery</b>	1.00	0.821
Stage 3	<b>Initial Investigation</b>	1.00	0.820
Stage 4	<b>Uncovering Secrets and Clues</b>	1.00	0.920
Stage 5	<b>Confrontation of Suspects</b>	1.00	0.798
Stage 6	<b>Red Herrings and False Leads</b>	1.00	0.925
Stage 7	<b>Realization or Insight</b>	1.00	0.797
Stage 8	<b>Final Confrontation</b>	1.00	0.842
Stage 9	<b>Resolution and Aftermath</b>	1.00	0.872

### Romance

Stage	Concept	Rate	Imp.
Stage 2	<b>Meeting with Mentor</b>	1.00	0.898
Stage 5	<b>Approach to Inmost Cave</b>	1.00	0.871
Stage 7	<b>Reward</b>	1.00	0.784
Stage 8	<b>Return</b>	1.00	0.809
Stage 9	<b>Resurrection</b>	1.00	0.639
Stage 10	<b>Return with Elixir</b>	1.00	0.637

### Satire

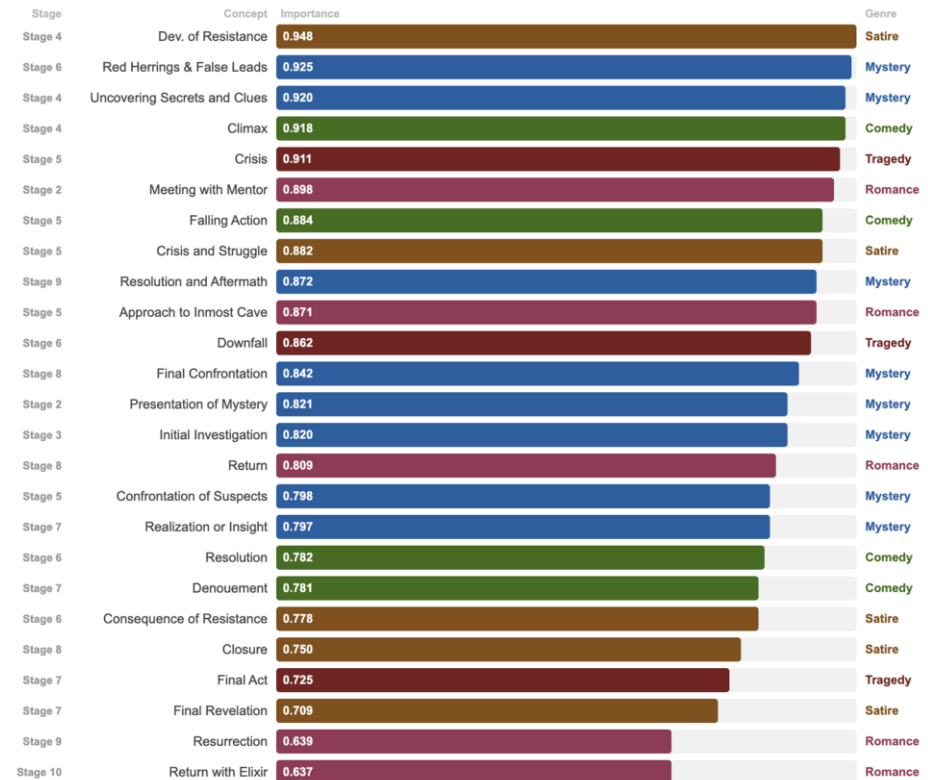
Stage	Concept	Rate	Imp.
Stage 4	<b>Development of Resistance</b>	1.00	0.948
Stage 5	<b>Crisis and Struggle</b>	1.00	0.882
Stage 6	<b>Consequence of Resistance</b>	1.00	0.778
Stage 7	<b>Final Revelation</b>	1.00	0.709
Stage 8	<b>Closure</b>	1.00	0.750

### Tragedy

Stage	Concept	Rate	Imp.
Stage 5	<b>Crisis</b>	1.00	0.911
Stage 6	<b>Downfall</b>	1.00	0.862
Stage 7	<b>Final Act</b>	1.00	0.725

Selection rate = 1.0 across all 200 subsampling runs

■ Mystery ■ Romance ■ Satire ■ Comedy ■ Tragedy



# Selection of Key Narrative Elements

## The least genre-distinctive structural elements

Stage	Concept	Rate	Imp.
Stage 1	<b>Set Up</b>	0.00	0.000
Stage 2	<b>Catalyst</b>	0.00	0.000
Stage 3	<b>Rising Action</b>	0.00	0.000

Stage	Concept	Rate	Imp.
Stage 1	<b>Introduction of Characters</b>	0.73	0.601

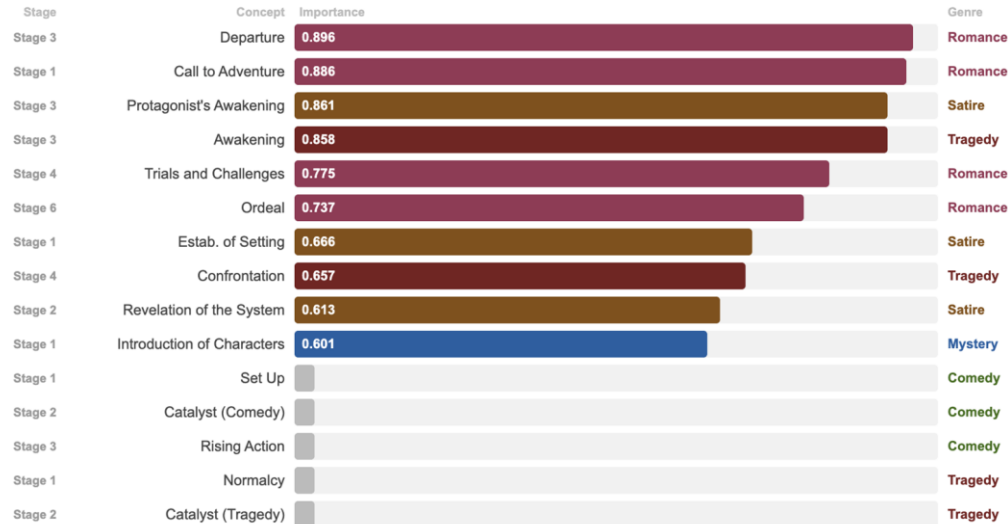
Stage	Concept	Rate	Imp.
Stage 4	<b>Trials and Challenges</b>	0.95	0.775
Stage 6	<b>Ordeal</b>	0.98	0.737

Stage	Concept	Rate	Imp.
Stage 1	<b>Establishment of Setting</b>	0.75	0.666
Stage 2	<b>Revelation of the System</b>	0.90	0.613
Stage 3	<b>Protagonist's Awakening</b>	0.97	0.861

Stage	Concept	Rate	Imp.
Stage 1	<b>Normalcy</b>	0.00	0.000
Stage 2	<b>Catalyst</b>	0.00	0.000
Stage 3	<b>Awakening</b>	0.95	0.858
Stage 4	<b>Confrontation</b>	0.75	0.657

Selection rate < 1.0 — structurally unstable or generic across genres

■ Mystery ■ Romance ■ Satire ■ Comedy ■ Tragedy



# Key Findings

- Developing a methodology for evaluating a literary model
- An LLM can aid in annotating narrative patterns, yet, it must be tuned and evaluated by human experts.
- The elements of the PatternTeller patterns are not genre-specific.
- Closure, Final Confrontation, and Return are the most common elements across all genres.
- The most genre-distinctive structural elements: *Development of Resistance* (Satire), *Downfall* (Tragedy), *Uncovering Secrets and Clues* (Mystery), etc.





Questions, please

**We welcome cooperation**  
**[milkak@post.bgu.ac.il](mailto:milkak@post.bgu.ac.il)**