Chair of Scientific Computing
Univ.-Prof. Dr. Hans-Joachim Bungartz

Bachelor Thesis Final Presentation

Exploring Fuzzy Tuning Technique for Molecular Dynamics Simulations in AutoPas

Manuel Lerchner
manuel.lerchner@tum.de

Advisors:
Manish Kumar Mishra, M.Sc. (hons)
Samuel James Newcome, M.Sc.

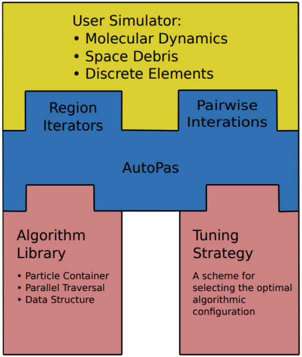
AutoPas

What is AutoPas?

- Library for optimal node-level performance in N-body simulations
- Many different implementations for the N-body problem
- AutoTuning: Automatically switch between implementations
  - **Container:** How to find neighboring particles?
  - **Traversal:** How to handle multi-threading?
  - **Data Layout:** How to store particles in memory?
  - **Newton 3:** Can we exploit Newton's 3rd law?
  - ...
- Example applications:
  - `md_flexible` (Molecular Dynamics)
  - `sph` (Smoothed Particle Hydrodynamics)

Structure of AutoPas

- Three main components:
  - User Application
  - Algorithm Library
  - Tuning Strategies
- Algorithm Library:
  - Huge Search Space[1]
- Tuning Strategies:
  - Full Search
  - Random Search
  - Predictive Tuning
  - Bayesian Search
  - Rule Based Tuning

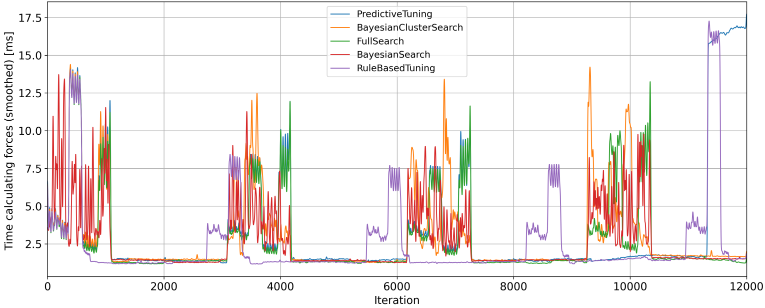


Source: [Newcome et al., 2023]

[1] Container × Traversal × Data Layout × Newton 3 × Load Estimator × Cell Size Factor

Auto-Tuning

- Tuning Phase: Find the best configuration
  - Tuning Strategies select configurations to evaluate
  - Expensive, Time consuming
- Simulation Phase: Use the best configuration



Fuzzy Logic Systems

- Use human-like reasoning to model complex systems
- Example: Heater Control
  - Input: temperature (e.g. 20°C), humidity (e.g. 50%)
  - Output: heater power (e.g. 50%)
  - Rules:

IF temp is *cold*	**AND**	humidity is *dry*	**THEN** power is *high*
IF temp is *hot*	**OR**	humidity is *wet*	**THEN** power is *low*
IF temp is *warm*			**THEN** power is *medium*

- Very easy to understand and interpret
- Can handle uncertainty and imprecise information
- Complexity is abstracted away in the linguistic terms (e.g. *cold, warm, hot*)
- Fuzzy System $f : \mathbb{R}^n \to \mathbb{R}$

Mathematical Foundations

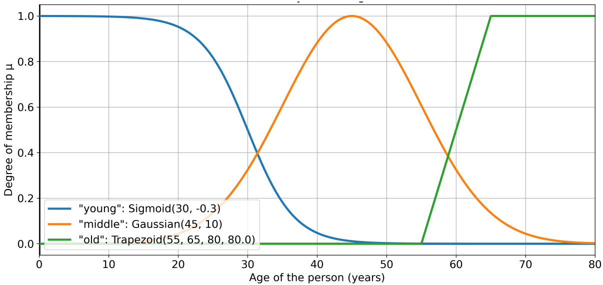
- Consider the Fuzzy Rule:

$$\underbrace{\underbrace{\textbf{IF} \; (\underbrace{\text{temp}}_{\text{Ling. Var.}} \; \text{is} \; \underbrace{\textit{cold}}_{\text{Ling. Term}} \; \textbf{AND} \; \underbrace{\text{hum}}_{\text{Ling. Var.}} \; \text{is} \; \underbrace{\textit{dry}}_{\text{Ling. Term}} \;)}_{\text{Antecedent}} \; \textbf{THEN} \; \underbrace{\underbrace{\text{power}}_{\text{Ling. Var.}} \; \text{is} \; \underbrace{\textit{high}}_{\text{Ling. Term}}}_{\text{Consequent}}}_{\text{Fuzzy Rule}}$$

- Fuzzy Logic Systems consist of:
  - Linguistic Terms / Fuzzy Sets (e.g. *cold, warm, hot*)
  - Linguistic Variables (e.g. temperature, humidity, power)
  - Fuzzy Logic Operators (e.g. **AND, OR, NOT**)
  - Fuzzy Rules (e.g. **IF** *antecedent* **THEN** *consequent*)

Fuzzy Sets

- Fuzzy Sets are generalizations of classical sets
  - Classical Sets: binary membership function $\in_A$: $A \rightarrow \{\mathit{false}, \mathit{true}\}$
- Fuzzy Sets are defined by:
  - Underlying Crisp Set X (e.g. *age* $\subseteq \mathbb{R}$)
  - **Continuous** membership function $\mu_{\tilde{A}} : X \rightarrow [0, 1]$
- Allow for uncertainty. When is a person young?



Linguistic Variables

- Linguistic Variables can take on linguistic terms / fuzzy sets
  - E.g. *age* can take *young*, *middle-aged* or *old*
- Instead of using crisp values (35 years), we use a combination of linguistic terms to describe the age (Fuzzyification):

$$\text{35 years} \implies \begin{cases} \text{20\% young} \\ \text{60\% middle-aged} \\ \text{0\% old} \end{cases}$$

- This allows for non-numerical reasoning:
  - E.g. **IF** age is *young* **THEN** fitness is *high*
  - Use abstract concepts instead of crisp values
- Each linguistic term represents a certain *collection* of values

Fuzzy Logic Operators

- Fuzzy Logic Operators are used to modify/combine fuzzy sets
- Extension of boolean logic operators to real numbers
  - $\wedge : \{\mathit{false}, \mathit{true}\} \times \{\mathit{false}, \mathit{true}\} \rightarrow \{\mathit{false}, \mathit{true}\}$
  - **AND** : $[0, 1] \times [0, 1] \rightarrow [0, 1]$
- Extended operators need to maintain the classical semantics
- Typically, Fuzzy Logic Operators are defined as:
  - **AND:** Corresponds to the intersection of fuzzy sets

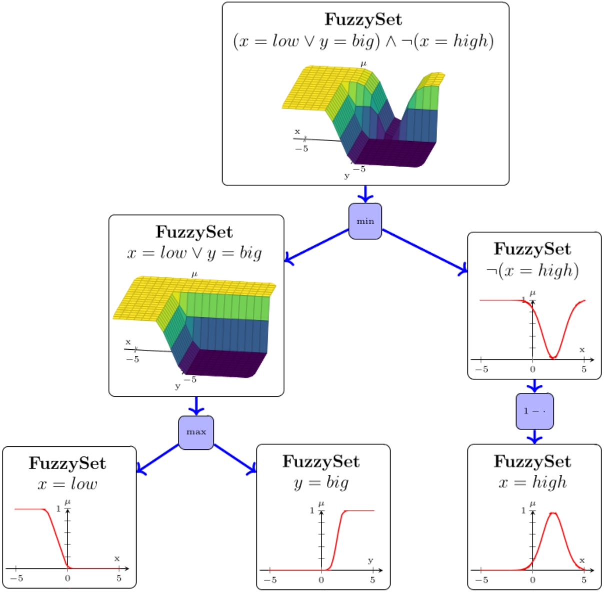
$$\mu_{\tilde{A}\cap\tilde{B}}(x) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$$

  - **OR:** Corresponds to the union of fuzzy sets

$$\mu_{\tilde{A}\cup\tilde{B}}(x) = \max(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$$

  - **NOT:** Corresponds to the complement of a fuzzy set

$$\mu_{\neg\tilde{A}}(x) = 1 - \mu_{\tilde{A}}(x)$$


FuzzySet
$(x = low \lor y = big) \land \neg(x = high)$
min
FuzzySet
$x = low \lor y = big$
FuzzySet
$\neg(x = high)$
max
$1 - \cdot$
FuzzySet
$x = low$
FuzzySet
$y = big$
FuzzySet
$x = high$

Fuzzy Rules

- Each rule is of the form: **IF** *antecedent* **THEN** *consequent*
  - Both antecedent and consequent are fuzzy sets
  - E.g. **IF** age is *young* **THEN** fitness is *high*
- Rules can be interpreted as a logical implication
  - **IF** $\tilde{A}$ **THEN** $\tilde{B}$ $\iff$ $\tilde{A}$ **IMPLIES** $\tilde{B}$
  - Implication is similar to previous operators (**AND, OR, NOT**)
  - Special form of implication: Mamdani Implication
    - $\tilde{R} =$ **IF** $\tilde{A}$ **THEN** $\tilde{B}$
    - $\mu_{\tilde{R}}(x) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$
  - *Effect* of the rule is limited by the *strength* of the antecedent
- A Fuzzy System can consist of multiple rules acting on the same linguistic variable
  - The total *effect* on the output is the combination/union of all individual rule outputs

Defuzzification

- Process of converting arbitrary fuzzy sets to a crisp value
  - Special case: Fuzzy sets resulting from rule application

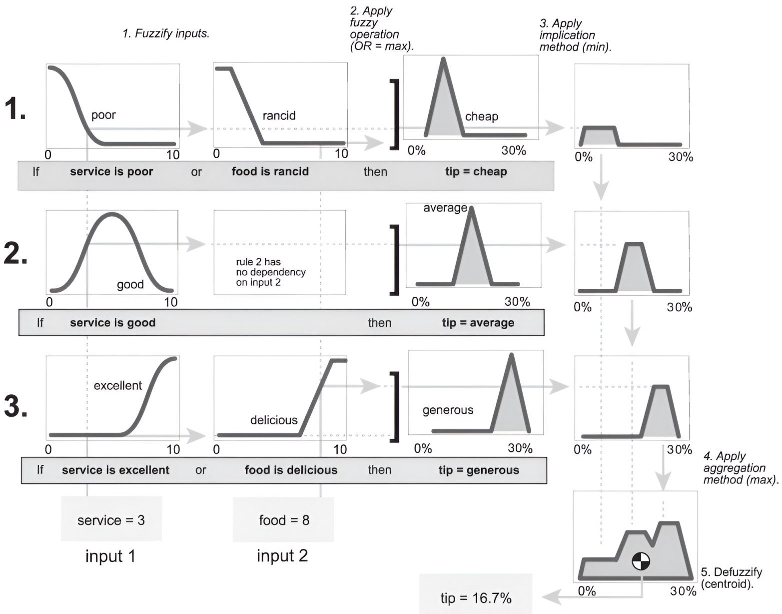
$$\begin{cases} 20\% \text{ young} \\ 60\% \text{ middle-aged} \\ 0\% \text{ old} \end{cases} \Longrightarrow 35 \text{ years}$$

- Multiple methods for defuzzification:
  - **Centroid:** Weighted average of the fuzzy set

$$\text{Centroid} = \frac{\int x \cdot \mu_{\tilde{R}}(x)dx}{\int \mu_{\tilde{R}}(x)dx}$$

  - **Mean of Maxima:** Average of all maxima of the fuzzy set
- Core idea: Represent aspects of the fuzzy set with a crisp value
  - E.g. Weighted average of all possible values (Centroid)
  - E.g. Most likely value (Mean of Maxima)

Source: MathWorks - Fuzzy Inference Process



- TODO: Maybe add decision surfaces here

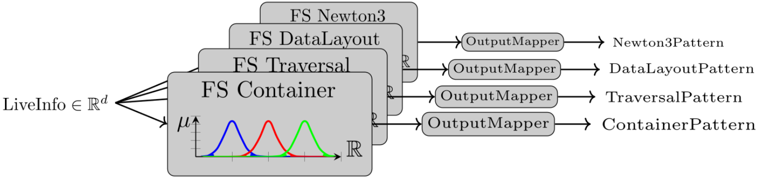
Fuzzy Tuning Strategy

- Main Idea: Use Fuzzy Logic to tune AutoPas
- Make use of LiveInfoData[2] to perform tuning
- Benefits:
  - Similar to Rule-Based Tuning
  - Potentially more expressive and powerful
  - Still easy to understand and interpret
- Challenges and Questions for AutoPas:
  - What are the output variables? How to predict Configurations?
  - How to interpret the result? (Fuzzy System : $f : \mathbb{R}^n \rightarrow \mathbb{R}$)
  - How to create the fuzzy rules? Expert knowledge?
  - How to specify the linguistic terms / fuzzy sets?

[2]Simulation state: `avgParticles/Cell, homogeneity, threadCount ...`

Approach 1: Component Tuning

- Independently predict good values/patterns for each tunable parameter
- Create a Fuzzy System for each tunable parameter
  - Container_DataLayout
  - Traversal
  - Newton 3
- Output Variables: Nominal Representation of parameter values
- Numeric output mapped to the closest value [Mohammed et al., 2022]
- Combine the patterns to obtain final configurations

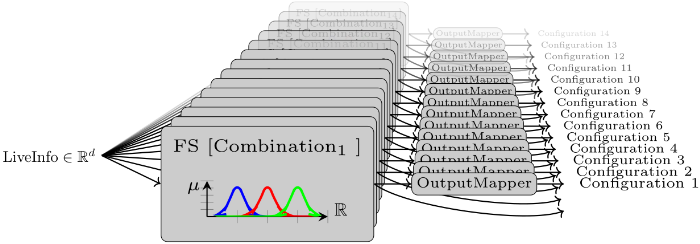


Approach 1: Component Tuning

- TODO Linguistic Variables

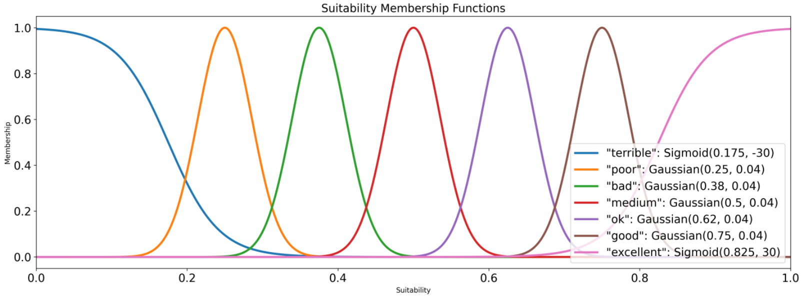
Approach 2: Suitability Tuning

- Predict the suitability of *each* configuration
- Use the suitability values to determine worthwhile configurations
- Create a Fuzzy System for each possible configuration
- Output Variables: Suitability of the configuration
- More complex, but potentially more powerful



Approach 2: Suitability Tuning

- Each configuration has a separate suitability variable
- Example rule:
  - **IF** threadCount is *high* **AND** avgParticlesPerCell is *low*
    THEN suitability_LinkedCells_AoS_lc_c18_disabled is *bad*
  - Where *high, low, bad* are appropriate linguistic terms



Data-Driven Rule Extraction

- Creating the Fuzy Rules is hard
  - Expert Knowledge is required
  - Formalization of knowledge is difficult
  - Potentially many rules required
- Use Machine Learning to extract rules from data
- Decision Tree $\rightarrow$ Fuzzy Decision Tree $\rightarrow$ Fuzzy Rules
  [Crockett et al., 2006]
- Does not require expert knowledge
- Human experts can still validate the rules

Conversion Process

Comparison and Evaluation

- Exploding Liquid Benchmark
- Spinodal Decomposition MPI
- Further Analysis

Future Work

- Dynamic Rule Generation
- Improving Tuning Strategies
- Simplification of the Fuzzy System

Conclusion

- Summary of Findings
- Impact
- Final Thoughts

References I

Crockett, K., Bandar, Z., Mclean, D., and OShea, J. (2006).
On constructing a fuzzy inference framework using crisp decision trees.
Fuzzy Sets and Systems, 157(21):2809–2832.

Mohammed, A., KorndÃ¶rfer, J. H. M., Eleliemy, A., and Ciorba, F. M. (2022).
Automated scheduling algorithm selection and chunk parameter calculation in openmp.
IEEE Transactions on Parallel and Distributed Systems, 33(12):4383–4394.

References II

Newcome, S. J., Gratl, F. A., Neumann, P., and Bungartz, H.-J. (2023).
Towards the smarter tuning of molecular dynamics simulations.
Amsterdam, The Netherlands.