

# Making computer science research more accessible in India

January 31 2023, by Alex Shipp



Credit: CSAIL/MIT

Imagine that you are teaching a technical subject to children in a small village. They are eager to learn, but you face a problem: There are few resources to educate them in their mother tongue.

This is a [common experience](#) in India, where the quality of textbooks written in many local languages pales in comparison to those written in English. To address educational inequality, the Indian government launched an [initiative](#) in 2020 that would improve the quality of these resources for hundreds of millions of people, but its implementation remains a massive undertaking.

Siddhartha Jayanti, an MIT Ph.D. student in electrical engineering and [computer science](#) (EECS) who is an affiliate of MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) and Google Research, encountered this problem first-hand when teaching students in India about math, science, and English. During the summer after his first year as an undergraduate at Princeton University, Jayanti visited the town of Bhimavaram, volunteering as an organizer, teacher, and mentor at a five-week education camp. He worked with economically disadvantaged children from villages across the region. They spoke Telugu, Jayanti's mother tongue, but faced linguistic barriers because of the complex English used in [academic work](#).

According to the [World Economic Forum](#) and U.S. Census data, Telugu is the United States' fastest-growing language, while [Ethnologue](#) estimates over 95 million speakers worldwide, further emphasizing the need for more academic materials in the vernacular.

As a distributed computing and AI researcher with a shared cultural background, Jayanti was in a unique position to help. With millions of Telugu speakers in mind, Jayanti wrote the first original computer science paper to be composed entirely in Telugu in 2018. This research then became publicly accessible on *arXiv* in 2022, focusing on designing simple, fast, scalable, and reliable multiprocessor algorithms and analyzing fundamental communication and coordination tasks between processors.

Processors are electronic circuitry that execute computer programs, making them notorious for their many moving parts. "Think about processors as people completing a task," says Jayanti. "If you have one processor, that is like one person doing a task. If you have 200 people instead, then ideally your team will solve problems faster, but this is not always the case. Coordinating multiple processors to achieve speedups requires clever algorithmic design, and there are sometimes fundamental communication barriers that limit how fast we can solve problems."

To solve computing problems, each process in a multicore system follows a strict procedure, which is also known as a multiprocessor algorithm. Still, there are certain limits on how quickly processors can interact with each other to compute solutions. Jayanti's paper highlighted a key communication bottleneck for these algorithms, known as generalized wake-up (GWU), where a processor "wakes up" when it has executed its first line of code.

But the question remains: Can each processor figure out that the others have woken up? Jayanti indicates that the answer is yes, but due to the work each solution requires, there are certain mathematical limits to how quickly GWU can be resolved.

The issue is part of a larger trend: The multicore revolution, where many chip manufacturers are no longer prioritizing faster processing speed. Instead, chips are now commonly designed with multiple cores, or smaller processors within larger CPUs. Multicore chips are now commonplace in many phones and laptops.

"Modern technology requires simple, fast, and reliable multiprocessor algorithms," says Jayanti. "Huge speedups and better coordination is the goal, but even using multiprocessor algorithms, we can prove that communication problems can only be solved so quickly."

Overcoming significant linguistic barriers to communicating state-of-the-art research in Telugu, Jayanti invented new technical vocabulary for the paper using Sanskrit, the classical language of India, which heavily influences Telugu. For example, there was no word for technical terms like "shared-memory multiprocessor" in Telugu. Jayanti changed that, coining the word *saṁvibhakta-smṛti bahusaṁsādhakamu* (సామవిభక్తా-స్మృతి బహుసాంసాధకము).

While the term may seem daunting and complex at first, Jayanti's process was simple: Use Sanskrit root words to coin new words in Telugu. For instance, the Sanskrit root "vibhaj" means "to partition" while "smṛ" means "to remember, recollect, or memorize." After modifying these words with prefixes and suffixes, the results are "saṁvibhakta" ("shared") and "smṛti" ("memory"), or "saṁvibhakta-smṛti" ("shared-memory") in Telugu.

Passionate about creating [educational opportunities](#) in India, Jayanti has visited schools in several states, including Telangana, Andhra Pradesh, and Karnataka. He travels to India yearly, occasionally making stops at universities like the International Center for Theoretical Sciences and those within the Indian Institutes of Technology.

By creating new technical vocabulary, Jayanti sees his work as an opportunity to empower more people to pursue their dreams in science. His Telugu paper opens the doors for millions of native speakers to access STEM research.

"Knowledge is universal, brings joy, opens doors to new opportunities, and has the power to enlighten and bring people of diverse backgrounds closer together in pursuit of a better world," says Jayanti. "My scientific learnings and discoveries have brought me in contact with great minds around the world, and I hope that some of my work can open up a gateway for more people worldwide."

As part of his Ph.D. thesis, Jayanti proposed the Samskrtam Technical Lexicon Project, which would bridge further education gaps by developing a dictionary of modern technical terms in STEM for speakers of local Indian languages and academics. "The project aims to forge a close collaboration between scholars of STEM, Sanskrit, and other vernaculars to expand science-availability in language communities that span over a billion people," according to Jayanti.

Jayanti's research also fueled further studies of multicore processing speeds. In 2019, he teamed up with Robert Tarjan, a professor of computer science at Princeton and Turing Award winner, as well as Enric Boix-Adserà, an MIT Ph.D. student in EECS to [demonstrate lower bound speed limits](#) for data structures like union-find, where algorithms can create a "union" between disjointed datasets while "finding" whether two items are currently in the same set.

The team leveraged Jayanti's research on GWU to prove certain limits on how fast algorithms can be, even harnessing the power of multiple cores. Jayanti and Tarjan have designed some of the fastest algorithms for the concurrent union-find problem yet, making analysis of large graphs like the internet and road networks much more efficient. In fact, these algorithms are close to the mathematical speed barrier for solving union-find.

Jayanti's 2018 research paper in Telugu was presented along with an abstract in Sanskrit as one of the 14 chapters of his thesis last year, and his team's 2019 paper was presented at the Symposium on Principles of Distributed Computing. His graduate studies were supported by the U.S. Department of Defense through the National Defense Science and Engineering Graduate Fellowship.

**More information:** Siddhartha Visveswara Jayanti, Generalized Wake-Up: Amortized Shared Memory Lower Bounds for Linearizable Data

Structures, *arXiv* (2022). [DOI: 10.48550/arxiv.2207.07561](https://doi.org/10.48550/arxiv.2207.07561)

Provided by MIT Computer Science & Artificial Intelligence Lab

Citation: Making computer science research more accessible in India (2023, January 31)  
retrieved 27 April 2024 from <https://techxplore.com/news/2023-01-science-accessible-india.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.