

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 2. ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ И СИСТЕМЫ
ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 6

Москва 2015

ОБЩИЙ РАЗДЕЛ

УДК 004.5 : 004.7

М.Н. Тихонов, М.М. Богословский

Факторы интернет-зависимости

Обсуждается (недостаточно освещаемая в отечественной прессе) крупная социально-психологическая проблема, обусловленная современным научно-техническим прогрессом в киберпространстве. Подробно рассматриваются негативные последствия патологического времяпрепровождения и характерные для интернет-зависимых людей (взрослых и детей) способы девиантного поведения в Сети, а также профилактика интернет-аддикции. Рассмотрены основные негативные факторы риска воздействия на здоровье человека, возникающие при продолжительной работе на компьютере. Затронута тема синдрома компьютерного стресса.

Ключевые слова: компьютер, современные сетевые технологии, интернет-аддикция, негативные аддиктивные активности, киберпреступления, дети онлайн, здоровье, риск, профилактика интернет-зависимости

ЧЕЛОВЕК В МИРЕ КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ

Современное производство, науку и технику невозможно представить без использования информационных технологий [1]. Казалось бы, совсем недав-

но слово «персональный компьютер» (ПК) вошло в нашу лексику, а сегодня мы уже не мыслим себя без этой техники. Вся относительно простая и сложная, рутинная работа в управлении производством, промышленностью и хозяйством передается автоматам и электронно-вычислительным машинам (ЭВМ).

В настоящее время трудно назвать такую область производственной или бытовой деятельности людей, в которой не использовались бы ЭВМ. Автоматизированные системы управления (АСУ), оперативные системы представления данных, информационно-справочные системы с неограниченно большим количеством справочно-библиографической информации, АСУ финансово-хозяйственной деятельности предприятий, социально-гигиенические и медико-экологические мониторинги, базы и банки данных и знаний, автоматические системы сбора и передачи данных, компьютерные сети – вот далеко не полный перечень реальных всемирно-исторической значимости ЭВМ и новых информационно-коммуникационных технологий. Компьютер стал привычным явлением не только в офисах и научных лабораториях, но в школах, больницах, домах.

Какую роль в современном мире играет процесс обмена информацией, объяснять никому не нужно. Крупные корпорации, имеющие филиалы в разных городах и странах, постоянно нуждаются в обеспечении эффективности передачи производственной информации – как между своими подразделениями, так и с партнерами и поставщиками.

Компьютерные технологии обогатили человечество множеством новых неожиданных возможностей. Возник целый фейерверк телекоммуникационных технологий, зачастую с непредсказуемыми последствиями для человечества.

По глубине и широте проникновения в другие отрасли знаний интернет-технологии занимают сегодня одно из ведущих мест в информационном мире. Постепенно охватывающая все стороны общественной жизни сеть Интернет, пользователями которого в мире на конец 2001 г. являлось свыше 530 млн человек, – одна из самых глобальных структур на Земле и наиболее яркое проявление глобализации и информатизации социально-экономических процессов [2].

С увеличением количества пользователей и доли интернет-услуг в мировой экономике Всемирная сеть начинает оказывать существенное влияние на сложившиеся территориально-общественные системы, включая как хозяйство, так и общество в целом.

Появление мировых открытых сетевых информационных технологий означает начало новой эры в развитии методов и средств переработки информации. В таких сетях пользователь воспринимает себя и воспринимается другими как неотъемлемый компонент единого мирового информационного сообщества. Это огромная победа человеческого разума.

Различают два вида обращения человека с ресурсами и возможностями, которые предоставляет Интернет. В первом случае Интернет рассматривается как мощный инструмент, расширяющий возможности человека, – как средство коммуникации и поиска различной информации, во втором – как условие воспроизводства определенного психологического состояния индивида.

Человечество создает и развивает новые информационные технологии, чтобы обеспечить контроль за происходящим, но в то же время эти технологии способны контролировать нас и управлять нами. Это взаимный процесс. Любой прогресс в науке или тех-

нике, наряду с ярко выраженными безусловно положительными явлениями, неизбежно влечёт за собой и отрицательные стороны. Вопросы компьютеризации общества сейчас стоят в ряду множества факторов, влияющих на здоровье людей. Именно поэтому так важно оценить степень влияния информационных технологий на здоровье человека.

ПРОБЛЕМАТИКА ИНТЕРНЕТ-ЗАВИСИМОСТИ

Волна компьютеромании с игровой двухмерной жизнью и оторванностью от настоящей реальности захлестнула Россию [3]. К концу 2002 г. в Российской Федерации компьютер использовали по различным источникам от 5 до 18 млн человек, и их количество увеличивалось ежегодно на 400-500 тыс. В России в 2002 г. пользователями Интернета являлись 8,7 млн человек старше 18 лет, а в 2008 г. – 34,4 млн. человек, т. е. более 10% населения. По числу пользователей от 16 лет и старше на 2008 г. Россия занимает второе место в Европе после Германии: в Германии – 46 млн. человек, в России – 38 млн человек.

Появление широкого доступа в Интернет повлекло за собой существенные изменения в жизни современного человека, обусловленные удобством использования Всемирной сети и широким спектром предлагаемых услуг. Нашей второй натурой стали самые разнообразные действия, выполняемые онлайн, например: делать покупки, знакомиться с новостями, резервировать билеты на самолеты или номера в гостиницах, выполнять научную работу. Получили быстрое развитие различные виртуальные сообщества – раз попав в них, люди оставались там жить.

Основными потребностями, реализующимися с помощью онлайн-игр, являются отдых и развлечения, эмоциональное равновесие, возбуждение и преодоление препятствий и уход из реальности. Онлайн-игры предоставляют пользователям возможность реализации неудовлетворенных желаний и потребностей и приобретения вещей, которых не хватает в реальной жизни. Это своеобразная попытка выйти за рамки будничной жизни с целью обрести иное пространство для выживания и ощущения безопасности [4].

Информационная цивилизация претерпевает информационную революцию с неизбежными негативными информационно-обусловленными эмоциональными стрессами, что является одним из существенных факторов риска возникновения ряда заболеваний [5].

Интернет-коммуникация предоставляет человеку широкие возможности в плане динамики потребности-мотивационной сферы (познавательной, развлекательной, коммерческой, игровой и даже противоправной, а также готовности к преодолению возникающих проблем). Однако за все получаемые преимущества и выгоды общество вынуждено платить. Эта плата представляется в виде новой, на данный момент только формирующейся, проблемы патологического использования Интернета, или интернет-зависимости [6]. В самом общем виде интернет-зависимость определяется как «нехимическая зависимость от пользования Интернетом». Её характеризует навязчивое желание выйти в Интернет, находясь в off-line, и неспособность выйти из сети, будучи on-line.

Аддикция определяется через поведение как систематическое повторение действий, сосредоточенных на узконаправленной сфере деятельности при игнорировании ведущих социальных потребностей, с уходом от реальности и формированием личности зависимого типа [7]. Пренебрежение домашними и служебными обязанностями, отказ от социального взаимодействия, изоляция, разрушение связей с внешним миром, общая аутизация и инфантилизация личности – далеко не полный список проблем, обусловленных чрезмерным увлечением Интернетом [8]. В технологическом плане – это аддиктивный потенциал аддиктивного агента. В случае компьютерной игровой зависимости – это система специфических стереотипно повторяющихся действий в компьютерной игре на ПК, позволяющих в иллюзорно-виртуальном ключе разрешать конфликты и избегать проблем.

Формирование интернет-зависимости обусловливается личностно-характерологическими особенностями человека, немаловажную роль при этом играют микро- и макросоциальные факторы [9].

В случае формирования интернет-зависимости очень важными дополнительными факторами выступают присущие Интернету свойства глобального средства коммуникации, системы доступа к информации, метаперсонифицирующие свойства, особенности сети как культурной, коммерческой и профессиональной среды [2, 3, 5].

Интернет – это пространство анонимности, свободы, вседозволенности. Именно анонимность и бестелесность делают Интернет привлекательной средой для проявления разнообразных форм девиантного, агрессивного поведения пользователей интернет-сети [8, 9].

В связи с информационным бумом образовалась новая форма загрязнения – информационное загрязнение. Средства массовой информации и телекоммуникации, особенно кино и телевидение, все более переполнены сценами изощренного насилия, пошлости и непристойности. Особую обеспокоенность общественности вызывает прогрессирующий рост киберпреступлений (мошенничество, сексуальная эксплуатация, вандализм, киберпреследование и другие проявления агрессии и враждебности) и терроризма.

Экстремистские и террористические организации, различные деструктивные культы активно осваивают Интернет. С помощью информационно-манипулятивных технологий они вовлекают в противоправную и преступную деятельность огромный процент молодых людей.

Интернет как основной элемент массовой культуры проповедует, учит, заставляет рассматривать агрессию, эксплуатацию и репрессии в качестве нормальных и естественных явлений, как само собой разумеющееся.

Современные тенденции состоят в том, что масштабы и состав киберпреступлений (интернет-вандализм, разработка и производство вирусов, мошенничество, интернет-воровство, киберпреследование, распространение детской порнографии и др.) вместе с технической изощренностью преступников будут резко увеличиваться, поскольку все больше

людей пользуются Интернетом и другими телекоммуникационными технологиями [4, 5, 9].

В плане виртуальности интернет-коммуникации следует рассматривать, прежде всего, как практически неограниченная возможность пользователей экспериментировать с собственным Я, создавая виртуальные личности и сетевые идентичности, превращаясь в процесс самоконструирования человека информационной эпохи. В зависимости о ценностных установок эта реальность может трактоваться либо как предоставляющая человеку компенсаторные возможности, либо как дающая ему принципиально новый опыт; при этом в обоих случаях сомнений в уникальности этих возможностей, как правило, не возникает.

Проблематика интернет-зависимости сегодня является традиционной в мире в серии публикаций, посвященных гуманитарным аспектам информационного общества [9]. В настоящее время практически ни одна научная публикация по проблеме «человек и новые информационные технологии» не обходится без упоминания проблемы интернет-аддикции. Несмотря на относительно недолгую (чуть больше десятилетия) историю своего изучения, обращает на себя внимание фактически лавинообразный рост исследовательского интереса к описываемой теме [9].

Зависимость от Интернета оказалась серьезной проблемой не только для США, но и для таких стран, как Китай, Корея и Тайвань, называемых «азиатскими тиграми», в которых привязанность достигла размеров поистине эпидемических. Зависимость от компьютера (психологическая невозможность совладать с желанием насытиться новой информацией) формируется намного быстрее, чем формирование традиционных аддикций (курение, алкоголизм, наркомания и игромания). Как указывает К.Янг, 25% интернет-аддиктов приобрели зависимость в течение полугода после начала работы в Интернете, 58% – в течение второго полугодия, а 17% – через год [10]. Опасность стать зависимым грозит каждому, кто проводит за видеоиграми более двух часов в день.

Будучи 15 лет назад предметом конкретных (носящих сугубо описательный характер) исследований в рамках психиатрии и психологии поведенческих зависимостей, сегодня эта проблематика широко представлена в социально-философских и культурологических, в социологических и социально-психологических работах, не говоря уже об общепсихологических. При этом вследствие более чем скромной истории изучения собственно научное знание об этой проблеме изобилует нерешенными вопросами, противоречивыми результатами, альтернативными моделями объяснения [9].

Научные исследования, касающиеся влияния интернет-зависимости на человека, в нашей стране пока очень редки, локальны и бессистемны. Эта территория психологической проблематики и психологических исследований.

Остается ряд неизученных вопросов, касающихся диагностических критериев такого расстройства, его нозологической принадлежности, особенностей личностных свойств интернет-зависимых и факторов риска его формирования. Не исследованы особенности клиники и течения данной формы зависимого по-

ведения среди подростков и юношей. Важнейшим является изучение влияния семейных отношений на формирование интернет-зависимости. Для нас эта проблема относительно нова.

В глобальном плане не вызывает сомнения актуальность и необходимость изучения зависимости между интернет-информацией и рядом типовых патологических процессов и нозологических форм. Нет сомнений ни в позитивном, ни в деструктивном воздействии средств массовой информации и, в первую очередь, Интернета, на поведение и психику человека, что находит всё больший общественный резонанс.

РИСКИ ДЛИТЕЛЬНОГО ВЗАИМОДЕЙСТВИЯ С СЕТЬЮ

Многие постоянные пользователи Интернета отмечают, что часто после длительной работы на компьютере появляются головная боль, болезненные ощущения в области мышц лица и шеи, ноющие боли в позвоночнике, резь в глазах, слезоточивость, нарушение чёткого видения, боли при движении рук. Часто присутствуют жалобы на онемение шеи, боль в плечах и пояснице или покалывание в ногах. Наиболее распространён кистевой туннельный синдром, при котором нервы руки повреждаются вследствие частой и длительной работы на компьютере. В наиболее тяжёлой форме этот синдром проявляется в виде мучительных болей, лишающих человека трудоспособности. Неподвижная напряжённая поза оператора, в течение длительного времени прикованного к экрану монитора, приводит к усталости и возникновению болей в позвоночнике, шее, плечевых суставах, а также развивается мышечная слабость и происходит изменение формы позвоночника. Интенсивная работа с клавиатурой вызывает болевые ощущения в локтевых суставах, предплечьях, запястьях, в кистях и пальцах рук [11].

Медикобиологические исследования воздействия компьютеров на пользователей показали, что степень болезненности ощущений пропорциональна времени работы за компьютером [12]. Негативное воздействие ПК на человека является комплексным, поэтому и изучение влияния компьютерных технологий должно быть системным, учитывающим взаимосвязанное комбинированное и сочетанное влияние множества факторов. Только комплексный подход позволяет достоверно оценить воздействие компьютера на здоровье пользователя [13-15].

Как показали результаты многочисленных научных работ, монитор компьютера является источником: электростатического поля; слабых электромагнитных излучений в низкочастотном и высокочастотном диапазонах (2 Гц, 400 кГц); рентгеновского, ультрафиолетового и инфракрасного излучений, а также излучений видимого диапазона [11,13-15]. Исследования функционального состояния пользователей ПК показали, что в организме человека под влиянием этих излучений происходят значительные изменения гормонального состояния, специфические изменения биотоков головного мозга, изменение обмена веществ. Низкочастотные электромагнитные поля при взаимодействии с другими отрицательными факторами могут

инициировать раковые заболевания и лейкемию. Пыль, притягиваемая электростатическим полем монитора, иногда становится причиной дерматитов лица, обострения астматических симптомов, раздражения слизистых оболочек [12]. Более 90% пользователей компьютеров жалуются на жжение или боли в области глаз, чувство песка под веками, затуманивание и неясность зрения, которая увеличивается в течение дня, и т.д. Комплекс этих и других характерных недугов с недавнего времени получил название «компьютерный зрительный синдром».

СПЕЦИФИЧЕСКОЕ ДЕЙСТВИЕ ЭЛЕКТРОМАГНИТНЫХ ИЗЛУЧЕНИЙ

Реакция человека на электромагнитное поле повышенной интенсивности проявляется, в первую очередь, поражениями иммунной, эндокринной и центральной нервной систем (ЦНС), выполняющих основные сигнально-регуляторные функции для поддержания организма в состоянии гомеостаза. Постоянные пользователи ПК чаще и в большей степени подвергаются психологическим стрессам, функциональным нарушениям и болезням сердечно-сосудистой системы [13].

Сроки возникновения нарушений в организме при облучении ЭМП зависят от многих факторов: частотного диапазона, продолжительности воздействия (стажа работы), локализации облучения (общее или местное), характера ЭМП (модулированное, непрерывное, прерывистое) и других.

Симптомы заболевания разнообразны и многочисленны [14,15]. Как правило, наличие единственного симптома маловероятно, поскольку все функциональные органы человека взаимосвязаны. При этом существенную роль играют индивидуальные особенности организма. Специалисты-медики выявили новый тип заболевания – синдром компьютерного стресса. Физические недомогания: сонливость, непреходящая усталость; головные боли после работы; боли в нижней части спины, в ногах; чувство покалывания, онемения, боли в руках; напряжённость мышц верхней части туловища.

Особый интерес представляют работы, касающиеся изучения влияния на центральную нервную систему низкоинтенсивных СВЧ-полей, модулированных в частотном диапазоне собственных биологических ритмов биообъекта. Установлено, что пороговые интенсивности для микроволновых излучений, модулированных в этом диапазоне, значительно ниже тех, которые являются характерными для импульсных и непрерывных излучений [10]. Низкоэнергетическое СВЧ-поле, модулированное в ритме собственных частот мозга, обладает выраженным кардиотропным действием.

При воздействии компьютера на мозговую (нервную) ткань с частотой, близкой к частоте собственных биоритмов мозга, биологическое воздействие увеличивается.

Особую значимость приобретает не интенсивность поля, а сам факт контакта с ним человека. Экспериментально установлено повышение чувствительности к ЭМП отдельных систем организма, вызванное другими факторами среды (физическими

полями, химическими и фармакологическими препаратами и т.п.). В этом случае даже при кратковременных контактах человека с ЭМП может возникнуть целый комплекс неврологических проявлений, психосоматических реакций, а также тяжёлых патологических реакций. Исходя из существующих теорий нетеплового механизма действия ЭМП (теория циклотронного резонанса, теория конформационных изменений и т.д.), возможны и другие проявления неблагоприятного воздействия полей на организм человека.

Пол, возраст и состояние алкогольного опьянения оказывают существенное влияние на чувствительность человека к слабым магнитным полям [16,17].

ВИРТУАЛЬНЫЙ МИР И РЕАЛЬНАЯ ОПАСНОСТЬ

С нашей точки зрения, весь современный спектр теоретических объяснений феномена интернет-аддикции в основе своей сводится к пониманию соотношения реальности и виртуальности.

Общеизвестно: все, что приносит удовольствие и служит способом удаления от рутинных текущих проблем, – потенциально обладает аддиктивными свойствами, т. е. может спровоцировать развитие зависимости. Интернет-среда предоставляет такую возможность в силу следующих своих характеристик [9]:

- неограниченный доступ к информации, уникальные возможности для коммуникации в различных виртуальных группах для самоутверждения собственного социального статуса;
- сверхличностная природа межличностных взаимоотношений и возможность анонимных социальных интеракций;
- вуайеристический аспект.

Контрастируя со «скучной и серой» реальностью, виртуальные взаимоотношения кажутся волнующими и многообещающими [4]. Интернет-зависимым свойственно мечтать, фантазировать, предвосхищать, предвкушать, чем они вскоре займутся в Интернете. Стремление к аддиктивной реализации обусловлено желанием избавиться от дискомфортного психологического состояния пустоты, скуки, подавленности, раздражения или беспокойства (неприятных эмоций и мыслей) [5]. Интернет-зависимые, находясь в интернет-среде, отмечают у себя эйфорию, ощущение «взлёта», возбуждение, усиление воображения, чувство свободы и беззаботности. Неадекватные формы поведения могли бы никогда не проявиться и остаться латентными, если бы не Интернет.

Реальность интернет-коммуникации в силу своих объективных, технологических особенностей, восполняет и компенсирует дискомфорт и дефициты современного человека, позволяя удовлетворять ряд потребностей, которые сегодня обусловлены социально-экономической реальностью. К таковым относятся: потребности в стимуляции и поиске новых ощущений (новизны), потребность в безопасности и соответствующие ей многочисленные варианты самоконструирования [4, 5, 8].

Разрушительный характер аддикции проявляется в том, что способ аддиктивной реализации из средств

ва постепенно превращается в цель, становится стилем (смыслом) жизни, в процессе которой человек оказывается в ловушке постоянного ухода от реальной действительности [18].

В группе, склонных к интернет-зависимости, преобладает агрессивный стиль, характеризующийся импульсивностью и резкостью поведения, демонстрацией собственного превосходства, а также повышенной тревожностью, уязвимостью и, как следствие, оборонительно-нападающей позицией, часто исключающей близость в общении [18].

Патологическое (от 20 до 50 ч в неделю) времяпрепровождение в интернет-сети причиняет ущерб физическому, психологическому, межличностному, семейному, экономическому и социальному статусу.

Различные авторы выделяют следующие психологические симптомы интернет-аддикции [19, 20]: модификация настроения; ощущение пустоты, депрессии, раздражения в отсутствии работы за компьютером; пренебрежение служебными обязанностями и недооценка социальных контактов; потеря контроля за собственным поведением; пренебрежение семьёй и друзьями; проблемы с работой (учебой); ложь работодателям и родным о своей деятельности.

В результате многолетней психотерапевтической работы с интернет-зависимыми пациентами специалисты выделяют следующие черты личности интернет-аддикта в качестве основных: инфантильность, низкая самооценка, эгоцентризм, низкая общительность, недостаточная ответственность, дефицитарная психика с хронической неудовлетворенностью собой, склонность к аффективным колебаниям [6, 21].

Зависимость ограничивает человека в его стремлении к достижению подлинной гармонии с собой, с другими людьми и с окружающим миром. Патологическое использование Интернета вызывает депрессивные расстройства, в том числе с суицидальными тенденциями, наносит ощутимый ущерб деятельности в реальной жизни. интернет-аддикция может легко перейти в другую зависимость, последствия от которой могут быть более пагубными. Более трех четвертей опрошенных из группы интернет-зависимых связаны со злоупотреблениями алкоголем и наркотиками [20].

ПОДРОСТКОВАЯ ИНТЕРНЕТ-ЗАВИСИМОСТЬ

Подростковый и юношеский возраст представляет собой один из наиболее существенных факторов риска по развитию расстройств поведения, связанных с формированием зависимости от Интернета [16, 21-23], являясь прямым следствием информатизации обучения и быта. ПК стал сегодня электронным средством, вошедшим в семейный быт, и подросток не может быть исключен из этого информационного поля даже при формировании зависимых отношений. Преподавание информатики сегодня входит в программу школьного образования, и подросток должен иметь навыки работы на ПК, должен систематически работать на нём.

К 2015-2020 гг., когда практически все российские населенные пункты будут подключены к Интернету (согласно федеральной программе «Электронная Россия»), ситуация в нашей стране будет

примерно такой же. Впрочем, можно смело предполагать, что в крупных российских городах, компьютеризация населения уже вполне соответствует мировым стандартам

Сталкиваясь с сопротивлением окружающей среды, стремясь к увеличению собственных возможностей, подросток убеждается, что трудности реального мира легко преодолеваются в виртуальном пространстве [19, 24]. Простым кликом «мыши», либо нажатием на клавишу портативного устройства, можно вступить в виртуальное общение с разными людьми, ощутить захватывающее возбуждение в азартной игре, восхищение от выявленных в Сети фактов и изображений, выброс адреналина от яростной игровой атаки и громкой победы над врагом, или, наконец, простое удовольствие от доступности всего этого. Ведь в Интернете, как говорится, есть всё, одна проблема – найти нужное, забывая в поисковую строку ключевые слова.

Ощущение огромных возможностей при минимальных физических действиях – сильное искушение. Используя уникальные свойства Интернета (глобальность, анонимность, вседозволенность, сервис чатов и блогов, возможности разнообразных ролевых игр), подростки впадают в Интернет-зависимость. В результате охваченности игровой зависимостью жизнедеятельность ребёнка становится доминантой игровой деятельностью [16, 24-26].

Интернет-зависимость представляет собой активное, включенное в определенную когнитивную деятельность, состояние. Онлайн-зависимость определяют как поведенческие, так и физиологические аспекты, часто они подкрепляют друг друга по принципу замкнутой петли. Технологические зависимости могут рассматриваться как разновидность поведенческих зависимостей, включая такие ключевые компоненты, как сверхценность, видоизменение настроения, увеличение толерантности, симптомы отмены, конфликт с окружающими и самим собой.

В молодом возрасте недостаточно развиты механизмы сдерживания импульсивного поведения, саморегуляции и контроля эмоциональной сферы. Коммуникативность становится особо привлекательной на фоне импульсивности и слабости Я, сверхценность ПК-игр и Интернета становится следствием общей личностной незрелости, инфантилизма группы подростков [19, 24].

Группа склонных к интернет-аддикции, существенно неоднородна с точки зрения приверженности к различным стратегиям копинга (ситуационно специфического поведения) в зависимости от конкретной трудности совладания – сферы реальных взаимоотношений или же сферы виртуальных коммуникаций. Склонные к интернет-аддикции, как правило, выбирают конструктивные стратегии совладания [21, 22].

Современный ребенок развивается в условиях хронической социальной депривации (недостаток общения в семье и вне её) [16, 23, 25]. Поэтому его следует целенаправленно готовить к эффективному взаимодействию с другими людьми (в том числе и со сверстниками) в условиях не только непосредственного контакта, но и электронно-коммуникативного опосредования.

КОРРЕКЦИЯ ЗАВИСИМОГО ПОВЕДЕНИЯ

Поскольку сегодня использование Интернета является неотъемлемой частью современной жизни, то обучение навыкам здорового взаимодействия с Интернетом особенно важно ещё в раннем возрасте, а в случаях с уже сформировавшейся аддикцией требуется коррекция зависимого поведения, возвращение клиентам контроля над собственной жизнью (прежде всего, над использованием ресурсов Сети). Необходимо тщательно исследовать те потребности, удовлетворение которых человек осуществляет посредством Интернета (потребность в общении, признании, достижениях, принадлежности, самореализации, поиске новых ощущений). Речь идет о скрытых личностных механизмах, которые лежат в основе формирования аддикции. Психотерапевтическая работа должна фокусироваться на том, каким образом человек может удовлетворить свои потребности в реальном, а не в виртуальном мире.

При наличии риска развития патологических расстройств следует проводить многоплановый комплексный анализ, опираясь на социально-психиатрический подход, в рамках следующих динамических отношений: социальная ситуация развития, индивидуально-личностные свойства и личностно-характерологические реакции (отклонения поведения и привычного жизненного стереотипа в подростковом периоде типа хобби-увлечений, определенные личностные особенности с учетом присущих подростково-юношескому этапу психологических черт возрастного психического развития).

Терапия должна быть направлена не на конкретную аддиктивную реализацию, а на зависимую личность. Психотерапевтический подход к зависимой личности предполагает продолжительную (не менее полугода) индивидуальную и/или групповую терапию. Основными точками приложения психотерапии являются низкая самооценка, плохая переносимость фрустраций и слабый контроль над импульсами.

Цели психотерапии должны включать: повышение самооценки и самоосведомленности, усиление контроля над импульсами, увеличение стабильности межличностных отношений, социальную адаптацию; анализ текущих копинг-стратегий и освоение новых, более эффективных стратегий совладания; обучение целеполаганию, а также заботе о себе, своем душевном и физическом здоровье [16, 21-23, 25].

Для работы с интернет-зависимыми используется когнитивно-поведенческий подход, основа которого – обучение пациента навыкам самоконтроля и саморегуляции, структурированию собственного времени, релаксационным техникам. Особую роль играет работа с семьей интернет-зависимого. Требуется изменение стереотипа поведения в семье, благожелательного семейного климата (комфорта). Степень семейного уважения к этико-нравственным ценностям и правилам в таких семьях ниже нормативной [27]. Изучение личностных свойств, особенностей воспитания и семейных коммуникаций как факторов риска формирования интернет-аддикции является актуальным для современной России. Полученные данные могут быть использованы при профилактической

работе с подростками с целью предотвращения негативных последствий пользования Интернетом [17].

Важно обучать близких интернет-зависимого эффективным стратегиям поведения, изменить семейную систему коммуникаций, что ускорит процесс социальной реабилитации пациента. Требуется поощрение вовлеченности интернет-аддиктов в повседневные дела других членов семьи, а также увеличение длительности их совместного времяпрепровождения.

В заключение следует сказать: поскольку прогрессирующее развитие технологической составляющей Интернет-коммуникации неминуемо влечет за собой содержательную вариативность патологических способов ее использования, то для лучшего изучения возникшей проблемы необходимо шире привлекать новые междисциплинарные методы и методики, расширять выборки респондентов [28-31].

В отечественной психологии признается необходимость разработки комплексного подхода к изучению феномена интернет-аддикции. Предлагается обсудить вопрос о создании Национального комитета по нехимическим аддикциям, который мог бы участвовать в разработке законодательных актов, препятствующих эпидемии распространенности данных патологических форм поведения, проводить экспертную оценку компьютерных игр, способствовать созданию и развитию профилактических программ.

Анализируя результаты масштабного исследования интернет-зависимости в мире, можно предположить, что Cyber Disorder (киберрасстройство) войдет в Международную классификацию болезней DSM-V.

СПИСОК ЛИТЕРАТУРЫ

- Игнатъев М.Б. Информационные технологии в микро-, нано- и оптоэлектронике: монография. – СПб.: ГУАП, 2008. – 200 с.
- Перфильев Ю.Ю. Пространственное распространение сети Интернет в России как процесс диффузии инноваций // Вестник Моск. ун-та. Сер.5. География. – 2003. – №2. – С. 30-36.
- Кастельс М., Киселева Э. Россия в информационную эпоху // Мир России. – 2001. – №1.
- Юревич А.В. Нравственное состояние современного Российского общества // СОЦИС. – 2009. – №10(306). – С.70-79.
- Еляков А.Д. Российское общество в информационном измерении // СОЦИС. – 2009. – №7(303). – С.85-94.
- Войскунский А.Е. Феномен зависимости от Интернета // Гуманитарные исследования в Интернете / под ред. А.Е. Войскунского. – М.: Можайск – Терра, 2000. – С.100-131.
- Griffiths M.D. Internet addiction: Internet fuels other addictions // Student British Medical Journal. – 1999. – Vol.7. – P. 428-429.
- Войскунский А.Е. Актуальные проблемы зависимости от Интернета // Психологический журнал. – 2004. – Т. 25, № 1. – С. 90-100.
- Интернет-зависимость: психологическая природа и динамика развития / ред.-сост. А.Е. Войскунский. – М.: Акрополь, 2009. – 279 с.
- Young K.S. Internet addiction: The emergence of a new clinical disorder // CyberPsychology @ Behavior. – 1998. – Vol. 3, № 1. – P. 237-244.
- Абдуллаева С.Г., Абдуллаева О.С. Воздействие компьютера на здоровье человека // Безопасность жизнедеятельности. – 2013. – № 5. – С. 8-9.
- Демирчоглян Г.Г. Компьютер и здоровье. – М.: Лукоморье, Новый Центр, 1997. – 256 с.
- Тихонов М.Н., Богословский М.М. Обратная сторона новых сетевых технологий // Сознание и физическая реальность. – 2012. – Т. 17, № 11. – С. 46-55.
- Довгуша В.В., Тихонов М.Н. Электромагнитный фактор – источник множества заболеваний // Медицина экстремальных ситуаций. – 1999. – № 1. – С. 5-10.
- Жураковская А.Л. Влияние компьютерных технологий на здоровье пользователя // Вестник Оренбургского государственного университета. – 2002. – №2.
- Янг К. Диагноз: Интернет-зависимость // Мир Интернет. – 2000. – № 2. – С.24-29.
- Уайнхолд Б., Уайнхолд Дж. Освобождение от созависимости. – М.: Независимая фирма «Класс», 2005.
- Бодров В.А. Информационный стресс. – М.: ПЕР СЭ, 2000.
- Чудова Н.В. Особенности образа «Я» «жителя» Интернета // Психологический журнал. – 2002. – Т. 23, № 1. – С. 113-117.
- Лисецкий К.С., Литягина Е.В. Психология негативных зависимостей. – Самара: Универс-групп, 2006.
- Пережогин Л.О. Интернет-аддикция в подростковой среде, 2007. – URL: [www.document] <http://banderus2.narod.ru/77187.html>
- Шайдулина А.Ф. Интернет-зависимость – новая форма аддиктивного поведения у подростков // Сб. материалов Конгресса по детской психиатрии. – М., 2001. – С. 71-72.
- Белинская Е.П. Человек в информационном мире // Социальная психология в современном мире. – М.: Аспект Пресс, 2002. – С.203-220.
- Кузнецова Ю.М., Чудова Н.В. Психология жителей Интернета. – М.: Изд-во УРСС, 2007.
- Керделлан К., Грезийон Г. Дети процессора. – Екатеринбург: У-Фактория, 2006.
- Бабаева Ю.Д., Войскунский А.Е. Одаренный ребенок за компьютером. – М.: Сканрус, 2003.
- Емельянова Е.В. Кризис в созависимых отношениях. – СПб.: Речь, 2004.
- Тихонов М.Н. Электромагнитная безопасность: взгляд в будущее. Проблемы комплексной

защиты организма пользователей при эксплуатации компьютерной техники // Экологическая экспертиза. – 2005. – № 3. – С. 9-47.

29. Степанова М. Как обеспечить безопасное общение с компьютером // Народное образование. – 2003. – № 2.
30. Морозов А.А. Экология человека, компьютерные технологии и безопасность оператора // Вестник экологического образования в России. – 2003. – № 1.
31. Ушаков И.Б. Оценка физических характеристик мониторов современных персональных компьютеров с позиции стандартов безопасности и характера деятельности // Безопасность жизнедеятельности. – 2002. – № 7.

Материал поступил в редакцию 17.02.15.

Сведения об авторах

ТИХОНОВ Михаил Николаевич – эксперт Межотраслевого экспертно-сертификационного, научно-технического и контрольного центра ядерной и радиационной безопасности (РЭСцентр), Санкт-Петербург; академик Международной академии наук экологии, безопасности человека и природы (МАНЭБ).
e-mail: M2Bog1@yandex.ru

БОГОСЛОВСКИЙ Михаил Михайлович – доктор биологических наук, нейрофизиолог, профессор Невского института управления и дизайна, Санкт-Петербург; академик МАНЭБ.
e-mail: M2Bog1@yandex.ru

Компьютерная интеллектуальная ДСМ-система для прогнозирования послеоперационных осложнений*

Описываются результаты работы компьютерной интеллектуальной ДСМ-системы для прогнозирования послеоперационных осложнений. Задача прогнозирования появления осложнений после операций на брюшной полости (рака желудка, толстого кишечника и т.д.) поставлена Российским онкологическим центром им. Н.Н. Блохина РАМН.

Ключевые слова: интеллектуальная система, ДСМ-метод, медицинская диагностика

В последнее десятилетие все более востребованы интеллектуальные системы поддержки принятия решений, ориентированные на медицину. Возрастающее количество данных об обследованиях больных приводит к росту объема информации, которую необходимо учитывать врачу при принятии решения о лечении. Наш опыт общения с экспертами-врачами показывает, что бывает трудно проанализировать множество результатов специальных обследований больного, и им хотелось бы иметь интеллектуальную систему, помогающую в принятии решения.

Среди многочисленных методов анализа данных ДСМ-метод автоматического порождения гипотез выделяется возможностью находить причинно-следственные связи на небольших выборках данных, когда статистические методы не работают. Это свойство позволяет эффективно использовать этот метод для медицинских прогнозов, где количество объектов исследований растет медленно.

В Отделе теоретических и прикладных проблем информатики ВИНТИ РАН создана интеллектуальная ДСМ-система (ИнтДСМ) [1]. ИнтДСМ является инструментом поддержки медицинских исследований со сложноструктурированными данными и множеством фактов, необозримых без использования компьютерных технологий, а также средством усиления интеллектуальных возможностей экспертов-медиков.

В настоящее время созданы ДСМ-системы для следующих медицинских задач.

1. Прогнозирование высокопатогенных типов вируса папилломы человека (ВПЧ) по цитологическим результатам исследования мазков (Кафедра клиниче-

ской и лабораторной диагностики Российской Медицинской Академии последипломного образования).

2. Диагностика двух заболеваний глаз: дегенеративного ретиношизиса и наследственных витреоретинальных дистрофий (Лаборатория клинической физиологии зрения МНИИ глазных болезней им. Гельмгольца).

3. Диагностика системной красной волчанки (Отделение нефрологии Городской клинической больницы им. Боткина).

4. Прогнозирование общей продолжительности жизни на основании анализа клинических и лабораторных данных (Российский онкологический центр им. Н.Н. Блохина РАМН).

5. Прогнозирование продолжительности жизни больных меланомой кожи и оценка прогностического значения уровня цитокинов и биомаркеров для прогрессирования меланомы кожи на фоне вакцинотерапии (Российский онкологический центр им. Н.Н. Блохина РАМН).

6. Прогнозирования результатов лечения и возможности рецидивов РМЖ (рак молочной железы) (Российский онкологический центр им. Н.Н. Блохина РАМН).

В настоящее время Российским онкологическим центром им. Н.Н. Блохина РАМН авторам была поставлена задача прогнозирования появления осложнений после операций на брюшной полости (рака желудка, толстого кишечника и т.д.).

Результатом работы является компьютерная интеллектуальная система прогнозирования клинического течения раннего послеоперационного периода у онкологических больных. Эта система устанавливает возможный риск развития гнойно-септических осложнений в раннем послеоперационном периоде у онкологических больных с новообразованиями торакoабдоминальной области IIa – IIIb стадии.

* Работа выполнена при поддержке программы фундаментальных исследований Президиума РАН (П15, проект №209) и гранта РФФИ (проект 14-07-00856а).

Область анализа клинических данных онкологических больных отвечает условиям применимости ДСМ-метода:

(1) возможность структурирования данных и формального определения сходства фактов из базы фактов (БФ);

(2) наличие положительных и отрицательных примеров ($\pm$)-примеров в БФ;

(3) наличие в БФ неявно заданных зависимостей причинно-следственного типа.

ИнтДСМ представляет собой компьютерную программу, работающую на персональном компьютере под управлением Windows 7, Windows 8. Для возможности создания ИнтДСМ для этой конкретной задачи необходима настройка на предметную область [2-4], включающая:

(1) разработка языка представления данных;

(2) определение понятия «объект» и «свойство» в терминологии ДСМ-метода;

(3) задание операции сходства;

(4) задание отношения вложения;

(5) определение аксиом предметной области.

Первым этапом совместной работы с врачами была разработка подсистемы представления знаний и создание БФ – одной из составных частей ИнтДСМ [1]. БФ представляет собой описания историй болезней на специальном формальном языке. Для ввода и отображения медицинских данных служит настраиваемый графический интерфейс. Настройка осуществляется с помощью специальных конфигурационных файлов, задающих формат и способ обработки клинических данных. В соответствии с языком представления данных внешний вид редактируемых параметров в интерфейсе программы максимально приближен к тому, что обычно используют врачи при заполнении карт.

Таким образом, задача врача – перечислить все возможные признаки описания больного, по которым надо сделать прогноз и предоставить достаточное количество примеров положительного и отрицательного типа (историй болезней больных с осложнениями и без осложнений). Задача разработчиков системы – структурировать эти признаки таким образом, чтобы была возможность применять к ним операцию сходства и определить отношение вложения. Эффективность анализа результатов различных исследований зависит от максимально полного описания обследований больного.

В БФ содержатся сведения о больном, включающие данные лабораторных исследований до и после операции, сопутствующие заболевания и т.д.¹ Цель системы – прогнозирование наличия хотя бы одного осложнения из некоторого списка (пневмония, панкреатит, перитонит, SIRS (наличие синдрома системной воспалительной реакции) и т.д.) и, по возможности, конкретизация осложнений.

База фактов создана на основе информации об онкологических пациентах, включая анамнез и результаты клинического исследования образцов крови, взятых:

до операции (за сутки до оперативного вмешательства); во время операции (полостной этап оперативного вмешательства); по окончании операции; в первые сутки после операции.

У пациентов, занесенных в БФ, анализировали следующие показатели крови: количество лейкоцитов ($\times 10^9$): относительное содержание нейтрофилов крови (%); фагоцитарный индекс (ФИ, %) – процент фагоцитирующих нейтрофилов в тест-системе латекс; ИЦА (Индекс цитотоксической активности, %) – по результатам МТТ-колориметрического теста. При соотношении клеток мишеней и клеток эффекторов 1:10; ЛПС (сывороточная концентрация липополисахарида, МЕ/мл); ИЛ-6 (сывороточная концентрация ИЛ-6, пг/мл); ИЛ-8 (сывороточная концентрация ИЛ-8, пг/мл); ИЛ-10 (сывороточная концентрация ИЛ-10, пг/мл); ИЛ-18 (сывороточная концентрация ИЛ-18, пг/мл); MCP-1 (сывороточная концентрация MCP-1, пг/мл); sCD14 (сывороточная концентрация sCD14, мкг/мл); разница НСТ-тестов (Разница показателей стимулированного латексом и базального НСТ-теста, для оценки кислородозависимого звена фагоцитоза) и т.д.

Из анамнеза в базу данных включена следующая информация: пол; возраст; диагноз; стадия опухолевого процесса (по классификации TNM); наличие сопутствующей патологии (сахарный диабет, открытый аортальный проток, хроническая почечная недостаточность, нарушение мозгового кровообращения, хроническая обструктивная болезнь легких, ишемическая болезнь сердца, хроническая анемия, гипертоническая болезнь, мерцательная аритмия, аортокоронарное шунтирование, постинфарктный кардиосклероз, атеросклероз); SOFA (оценка полиорганной недостаточности по шкале SOFA на первые сутки после операции) (баллы); SIRS - наличие синдрома системной воспалительной реакции; наличие послеоперационных осложнений (да/нет); вид послеоперационных осложнений (пневмония, панкреатит/панкреонекроз, эмпиема, сепсис, септический шок, перитонит, холангит, дистрибутивный шок).

(1) В описываемой ИнтДСМ для представления признаков, предоставленных врачами, используются два типа данных из семи возможных, перечисленных в [3, 4]:

Первый тип данных. Указывается один из возможных качественных признаков $E_i^{(1)} = At_i$, $i \in \{1, \dots, n\}$, n – число признаков в списке, например, признак «Фагоцитарный индекс до операции»: 0-40; 41-60; 61-100.

Второй тип данных. В кортеже длины « n » (где n – количество элементов списка признаков) указываются **присутствующие качественные признаки**, вместо отсутствующих признаков ставится Λ (пустой элемент):

$E_i^{(2)} = \langle At_1, \dots, At_i, \dots, At_n \rangle$, например, признак «Сопутствующие заболевания»: открытый аортальный проток, хроническая почечная недостаточность, нарушение мозгового кровообращения, хроническая

¹ Медицинские данные предоставлены аспирантами РОНЦ Плужниковой Н.А и Карпенко А.Ю.

анемия, мерцательная аритмия, постинфарктный кардиосклероз, атеросклероз, диабет и т.д.

(2) В рассматриваемой задаче объект, соответствующий истории болезни конкретного больного, представляет собой кортеж из 39 элементов:

$$O_i = \langle El_1, \dots, El_{39} \rangle$$

Каждый элемент этого кортежа соответствует конкретному признаку. В соответствии с языком представления данных, каждый элемент объекта принадлежит одному из двух типов данных п.(1).

Свойством является наличие какого-либо послеоперационного осложнения или же конкретного осложнения.

(3) Операция сходства определяется поэлементно.

$$\begin{aligned} O' &= \langle El'_1, El'_2, \dots, El'_{39} \rangle, \\ O'' &= \langle El''_1, El''_2, \dots, El''_{39} \rangle, \\ O' \cap O'' &= \langle El'_1 \cap El''_1, El'_2 \cap El''_2, \dots, El'_{39} \cap El''_{39} \rangle. \end{aligned}$$

Сходство элементов кортежей определяется в зависимости от того, какому типу данных принадлежит признак.

Для первого типа данных результатом операции сходства является совпадение признаков (в случае несовпадения – пустой элемент).

$$\begin{aligned} El^{(1)'} &= At_i \\ El^{(1)''} &= At_j \\ El^{(1)'} \cap El^{(1)''} &= \begin{cases} At_i, \text{ если } At_i = At_j \\ \Lambda \text{ иначе} \end{cases} \end{aligned}$$

Для второго типа данных результатом операции сходства является кортеж из признаков, присутствующих в обоих элементах признаков.

$$\begin{aligned} El^{(2)'} &= \langle At_{i_1}, \dots, At_{i_s}, \dots, At_{i_n} \rangle^2, \\ El^{(2)''} &= \langle At_{j_1}, \dots, At_{j_t}, \dots, At_{j_n} \rangle, \\ El^{(2)'} \cap El^{(2)''} &= \langle At_{i_1}, \dots, At_{i_s}, \dots, At_{i_n} \rangle, \\ \text{где } At_i &= \begin{cases} \Lambda, \text{ если } At_i' = \Lambda \text{ или } At_i'' = \Lambda \\ At_i' = At_i'' \text{ иначе} \end{cases} \end{aligned}$$

(4) Отношение вложения ($El \subseteq El'$) также определяется поэлементно и зависит от типа данных.

$$\begin{aligned} O' &= \langle El'_1, El'_2, \dots, El'_n \rangle, \\ O'' &= \langle El''_1, El''_2, \dots, El''_n \rangle, \\ O' \subseteq O'' &\Leftrightarrow El'_1 \subseteq El''_1 \& El'_2 \subseteq El''_2 \& \dots \& El'_{31} \subseteq El''_{31} \& \dots \& El'_n \subseteq El''_n. \end{aligned}$$

Для первого типа данных вложение равносильно совпадению значений атрибутов.

$$\begin{aligned} El^{(1)'} &= At_i, \quad El^{(1)''} = At_j, \\ El^{(1)'} \subseteq El^{(1)''} &\Leftrightarrow At_i = \Lambda \vee At_i = At_j. \end{aligned}$$

Для второго типа данных требуется присутствие в El' каждого из элементов El'' .

$$\begin{aligned} El^{(2)'} &= \langle At_{i_1}, \dots, At_{i_s}, \dots, At_{i_n} \rangle, \\ El^{(2)''} &= \langle At_{j_1}, \dots, At_{j_t}, \dots, At_{j_n} \rangle, \\ El^{(2)'} \subseteq El^{(2)''} &\Leftrightarrow \forall i (At_{i_1} = \Lambda \vee At_{i_1} = At_{j_1}). \end{aligned}$$

(5) Все перечисленные признаки, входящие в объект, можно разделить с учетом знаний о предметной области на три группы:

1) Группа необходимых признаков – признаки, без наличия которых гипотеза не имеет смысла;

2) Группа существенных признаков – признаки, без наличия хотя бы одного из которых, гипотеза не имеет смысла;

3) Группа сопутствующих признаков – признаки не входящие в группы 1) и 2), например, пол, возраст.

Таким образом, можно сформулировать следующую аксиому:

«В потенциальную гипотезу должны входить все признаки группы (1) и хотя бы один из признаков группы 2»).

Исходя из этой аксиомы предметной области, вводятся понятия фильтров. Конъюнктивный фильтр – требование вхождения в потенциальную гипотезу всех перечисленных в нем признаков. Дизъюнктивный фильтр – требование включения в потенциальную гипотезу хотя бы одного из перечисленных в дизъюнктивном фильтре признаков.

Настройка системы на экспериментальное исследование (препроцессинг) включает в себя следующее.

А. Выбираются стратегии: простой метод сходства, метод сходства с запретом на контрпримеры (невложение полученных методом сходства гипотез в исходные примеры противоположного знака) отдельно для (+) и (-)-примеров.

Б. Подбирается нужное количество «родителей» (минимальное количество примеров, формирующих гипотезу).

В. Подбирается конъюнктивный фильтр.

Г. Подбирается дизъюнктивный фильтр.

Возможный критерий оценки настройки системы на экспериментальное исследование – применение процедуры «доопределение по одному»: последовательно каждому объекту выборки присваивается значение «неизвестно», производится доопределение этого объекта средствами ДСМ-системы с выбранными параметрами и доопределенное значение сравнивается с существующим в БФ. Подсчитывается общее количество правильных и неправильных доопределений. Выбираются параметры пп.1-4, при которых при применении процедуры «доопределение по одному» будет наилучший результат, т.е. наибольшее количество правильных доопределений и наименьшее количество неправильных.

Для конкретного экспериментального исследования (22 больных с осложнениями, 29 – без осложнений) была произведена настройка, и выбраны следующие параметры:

1. Метод сходства с запретом на контрпримеры для положительных и для отрицательных случаев.

² Здесь и далее в объектах $i \in \{1, \dots, n\}$, n – число признаков в списке

2. Количество «родителей»: $K^+=10$, $K^-=10$
3. Конъюнктивный фильтр для «+» примеров – концентрация $sCD14^+$ до операции.
4. Дизъюнктивный фильтр – $CD14^+$ $CD16^+$ до операции или LPS во время операции.

В результате компьютерного исследования были получены следующие результаты: в 17-ти случаях – правильное положительное доопределение, в 24-х случаях правильное отрицательное доопределение, неправильных доопределений нет. Таким образом, точность данного исследования – 100%, а полнота – 83,7%.

Приведем пример прогноза. У пациента прогнозируется осложнение двумя положительными гипотезами: «Наличие у больного конкретного набора признаков есть причина появления осложнений»:

1.1) рак желудка	2.1) рак желудка
2) концентрация $IL-6$ до операции 0-30 мг/мл	2) концентрация LPS во время операции 0-0,1 МЕ/мл
3) концентрация $sCD14 + CD16^+$ до операции	3) концентрация $sCD14$ до операции
4) разность НСТ тестов до операции 0-0,6.	4) разность НСТ тестов до операции 0-0,6

В действительности этот прогноз оказался правильным: у больного развивалось осложнение SIRS.

Следующим этапом компьютерного исследования являлась конкретизация прогнозируемых осложнений. Возможность провести исследования определялась количеством положительных и отрицательных примеров. Достаточного количества примеров оказалось только для осложнения SIRS. С осложнением SIRS в БД оказалось 12 примеров, а без осложнения – 10).

Для настройки использовался метод с запретом на контрпримеры, количество родителей $K^+=6$, $K^-=4$, дизъюнктивный фильтр для отрицательного случая – $CD14^+$ $CD16^+$ до операции или LPS во время операции.

Порождено семь положительных и четыре отрицательные гипотезы. При применении процедуры «доопределении по-одному» получены следующие результаты: семь правильных и три неправильных доопределения. Точность доопределений – 70%, а полнота – 32%.

Так как было получено три неправильных доопределения, мы применили процедуру сокращения неинформативных признаков [5]. Было сокращено 11 неинформативных признаков и порождены две положительные и три отрицательные гипотезы.

Повторное применение «доопределения по-одному» дает следующие результаты: 10 правильных доопределений, неправильных нет. Точность доопределений – 100%, полнота 45,4%.

Приведем примеры правильных доопределений. У пациента прогнозировалось осложнение SIRS на основании гипотезы:

1. Диагноз – рак желудка.
2. Концентрация $IL-6$ до операции 0-30 мг/мл.
3. Концентрация $sCD14$ до операции 4,01 и более.
4. Разность НСТ-тестов до операции 0=0,6
5. Количество нейтрофилов до операции 50,01-70%
6. Концентрация LPS во время операции 0-0,1 МЕ/мл.

У пациента не прогнозировалось осложнение SIRS на основании гипотезы:

1. Стадия заболевания TNM_0 .
2. Концентрация LPS во время операции 0,11 МЕ/мл и более.

Таким образом, подтверждается возможность использования ДСМ-метода автоматического порождения гипотез в области медицинской диагностики, где недостаток формальных знаний может быть «скомпенсирован» богатым фактическим материалом. В некоторых областях медицины, находящихся на стадии феноменологического описания и накопления данных, развитые логико-математические методы восстанавливают причинно-следственные зависимости, служат для пополнения знаний на основе имеющихся фактических данных и являются средством поддержки интеллектуальных возможностей экспертов-медиков.

Описанная компьютерная система обладает высоким быстродействием, что позволяет проводить большой объем экспериментов по прогнозированию свойств в реальном времени.

Разработанная система используется в клинической практике РОНЦ им. Н.Н. Блохина РАМН.

Некоторые результаты работы системы использованы в диссертации на соискание ученой степени кандидата медицинских наук Плужниковой Натальи Андреевны «Иммунологические показатели как факторы прогноза развития гнойных осложнений и септиса у онкологических больных», 14.01.12, 14.03.09. – Москва, 2013.

Получено свидетельство о государственной регистрации программы для ЭВМ №201461187 «Компьютерная интеллектуальная система прогнозирования клинического течения раннего послеоперационного периода у онкологических больных».

* * *

Авторы выражают благодарность аспирантам РОНЦ Н.А. Плужниковой и А.Ю. Карпенко за предоставление медицинских данных.

СПИСОК ЛИТЕРАТУРЫ

1. Финн В.К. Эпистемологические основания ДСМ-метода автоматического порождения гипотез. Часть I // Научно-техническая информация. Сер.2. – 2013 – № 9. – С. 1-29.
2. Финн В.К. Эпистемологические основания ДСМ-метода автоматического порождения гипотез. Часть II // Научно-техническая информация. Сер.2. – 2013. – №12. – С. 1-26.

3. Панкратова Е.С. Интеллектуальная система типа ДСМ для анализа клинических данных // Научно-техническая информация. Сер.2. – 2011. – № 4. – С.8-16.
4. Панкратова Е.С., Виноградов Д.В. Формальное описание настройки интеллектуальных ДСМ-систем // Научно-техническая информация. Сер.2. – 2011. – № 9. – С.1-5.
5. Панкратова Е.С., Добрынин Д.А. Об одном способе выявления неинформативных признаков в интеллектуальных ДСМ-системах по медицине // Научно-техническая информация. Сер.2. – 2014. – № 3. – С. 10-14.

Сведения об авторах

ПАНКРАТОВА Елена Сергеевна – кандидат технических наук, старший научный сотрудник ВИНТИ РАН, Москва.
e-mail: pankr@viniti.ru

ДОБРЫНИН Дмитрий Анатольевич – кандидат технических наук, научный сотрудник ВИНТИ РАН, Москва.
e-mail: dobr@viniti.ru

Материал поступил в редакцию 03.04.15.

Сравнительный структурно-статистический анализ лексики и связей информационно-поисковых тезаурусов

Представлены результаты статистического анализа экспериментального объединенного информационно-поискового тезауруса, для построения которого использован теоретико-графовый подход. Приведены распределения, отражающие структурные различия тезаурусов, в том числе глубину детализации и наполненность понятий, типы связей, длину дескрипторов.

Предложено различать парадигматическую и синтагматическую полисемию. В случае синтагматической полисемии предложено указывать аспект при связях дескриптора. Определено понятие «семантической похожести» дескрипторов и приведены распределения, отражающие ее характер в зависимости от структурных параметров дескрипторов в различных тезаурусах.

Экспериментальные исследования проводились на массиве лингвистического ресурса, построенного путем объединения информационно-поисковых тезаурусов различных областей знаний.

Ключевые слова: тезаурус, дескриптор, структура связей, парадигматические отношения, полисемия

ВВЕДЕНИЕ

Основное назначение тезауруса, изначально определенного П.М. Роджетом как средство «для облегчения выражения мыслей и помощи при написании сочинений» (цит. по: [1]), практически не изменилось и разве что может быть дополнено явным указанием его коммуникативной функции: тезаурус в итоге должен обеспечить такие условия, чтобы мысли, зафиксированные посредством лексики тезауруса, были восприняты приемником адекватно авторскому представлению.

Роджетом тезаурус определялся как «обращенный словарь», т.е. «если в обычном словаре по слову находят его значение, то в тезаурусе по значению, которое записывается некоторым способом, находят слово или несколько слов, выражающих это искомое значение» (цит. по: [1]). Этот «некоторый» способ был очевидной иерархической структурой с шестью классами на верхнем уровне и тысячей тематических групп на нижнем.

Позднее и в первых определениях информационно-поискового тезауруса (ИПТ) отмечается, что тезаурус – это, прежде всего, классификационная система, в которой слова и выражения языка (конкретной предметной области) классифицируются по возможно ограниченному числу признаков [2]. При этом тезаурус как используемый человеком при индексировании и поиске физический инструмент

был представлен в виде книги и включал две основные части: 1) общий алфавитный список дескрипторных статей (заглавный дескриптор и связанные термины) и ключевых слов, снабженных пометами и ссылками; 2) упорядоченное множество озаглавленных тематических классов и групп, содержащих алфавитно-упорядоченные наборы дескрипторных статей.

В современных подходах тезаурус представляется в более общей форме – как семантическая сеть, в которой понятия связаны регулярными и устойчивыми семантическими отношениями – иерархическими (например, род-вид, целое-часть), ассоциативными, а также отношениями эквивалентности. При этом отдельное понятие определенной области знаний в тезаурусе представлено дескриптором (словом или словосочетанием), связанным с некоторыми другими терминами (дескрипторами или дескрипторами) так, что в совокупности они определяют место и роль этого понятия в данной предметной области. То есть понятие класса присутствует, однако, оно не является изначальным и определяющим, хотя и в стандарте [3], и в ряде промышленных тезаурусов дескрипторная статья сопровождается шифром предметной рубрики. Иерархические отношения в тезаурусе представляют собой классификацию, построенную на словах естественного языка, а не на абстрактных категориях, поэтому нарушается правильная структура дерева – один и тот же термин может иметь несколько «родителей» – вышестоящих терминов на преды-

душем уровне. Например, в Тезаурусе по информатике [4] словосочетание *Автоматизированная обработка информации* имеет два вышестоящих родителя: *Автоматизированная обработка* и *Обработка информации*, а слово *буквы* – целых три родителя: *алфавиты*, *символы*, *буквенно-цифровая информация*.

Такой отход от строгой упорядоченности обусловлен преимущественно практическими целями – повысить эффективность использования тезауруса в прикладных поисковых процессах путем включения на знаковом уровне свойств (объектов и отношений) собственно предметной области.

При этом, поскольку конкретные ИПТ традиционно создавались как специализированные – для отражения объектов и связей отдельной области знаний, информация о которых обычно была сосредоточена в соответствующей тематической базе данных, это, собственно, и позволяло локализовать смысловую неоднозначность знаковой системы и обеспечивать эффективность при индексировании и поиске.

1. ПОСТАНОВКА ЗАДАЧИ

Устойчивая тенденция создания и использования больших политематических баз данных приводит к необходимости создания политематического, а в пределе – всеобщего универсального тезауруса. Очевидно, что создаваться он может на базе отраслевых ИПТ, что вследствие тесной связи понятийно-знаковой системы тематического тезауруса с определенной отраслью знаний может негативно повлиять на эффективность его использования в разных предметных областях.

Исследования в области моделирования, анализа и использования ИПТ имеют давнюю историю и по-прежнему актуальны.

В [5] показано, что использование тематически значимых ключевых слов и семантических отношений между ними повышает эффективность поиска биологической информации на естественном языке.

В [6] рассмотрены вопросы, связанные с текстовым и геометрическим поиском географической и временной информации в негеографических информационных системах с использованием специального тезауруса.

Метод классификации патентных документов с использованием расширяемого тезауруса технических терминов описан в [7]. Авторы показали, что качество такой классификации выше, чем без применения специальной лексики.

В [8] для двух различных тезаурусов одной предметной области исследовался вопрос влияния объема тезауруса конкретной области знаний на качество сравнения схем в задачах интеграции данных.

В [9] представлены результаты эксперимента по внедрению в поисковую среду инструментов, позволяющих использовать при поиске тезаурус MeSH, что позволит устранить главное препятствие при поиске медицинской информации – незнание медицинской терминологии. Опыт использования тезауруса по медицине MeSH для автоматического расширения поискового запроса также приведен в [10, 11].

Результаты применения специализированного тезауруса при поиске решений в базе данных автосборочного завода Дженерал Моторс описаны в [12].

Анализ опыта Германской национальной экономической библиотеки по использованию тезауруса по экономике при индексировании, поиске и в процедурах автоматического расширения поисковых запросов представлен в [13].

Большой интерес вызывает использование таких лингвистических структур, как тезаурус WordNet при поиске в политематических ресурсах (например, в ресурсах Интернет). В этом направлении устойчиво применяются в основном только отношения для лексики естественного языка. Однако следует отметить, что такие лингвистические ресурсы отображают состав и структуру лексической системы естественного языка как такового, но не отражают особенности отдельных тематических областей или отраслей знаний. Соответственно, применение «общезыкового» ресурса в целях развития поисковых запросов может привести к понижению точности выдачи. Для поисковых машин Интернета, так или иначе использующих технологии расширения поисковых запросов, можно привести следующий пример. В ответ на запрос «продажа лука» уже первые страницы выдачи (ориентированной в основном на лук как растение) содержат ссылки на сайты оружия – арбалеты, луки, ножи, т.е. разрешить семантическую неоднозначность запроса только лексически не удастся.

В данной работе не будет рассматриваться эффективность тезаурусов непосредственно для повышения эффективности поиска или качества автоматической классификации и рубрицирования. Это достаточно убедительно и полно демонстрируют многие, в том числе и перечисленные выше, работы. Цель настоящей статьи – сравнительное статистическое исследование структурных компонентов лингвистического ресурса, построенного путем объединения нескольких тезаурусов, в контексте их принадлежности различным отраслям знаний.

2. МЕТОДИКА ФОРМИРОВАНИЯ ОБЪЕДИНЕННОГО ТЕЗАУРУСА

Для объединения тезаурусов различных областей знаний был применен теоретико-графовый подход.

Представим отдельный тезаурус помеченным взвешенным ориентированным графом $G_i = \langle V_i, X_i \rangle$, где $V_i = (v_1^i, v_2^i, \dots, v_{p_i}^i)$ – множество вершин, $p_i = |V_i|$; $X_i = (x_1^i, x_2^i, \dots, x_{q_i}^i)$ – множество дуг, $q_i = |X_i|$.

Каждая дуга задается тройкой: $x_k^i = (v_l^i, v_m^i, w_j)$, где v_l^i – вершина начала дуги; v_m^i – вершина завершения дуги; w_j – тип тезаурусной связи $w_j \in W$, где W – множество всех типов связей всех тезаурусов.

Тогда объединение тезаурусов можно представить как операцию объединения графов:

$$G_{\cup} = \bigcup_{i=1}^N G_i,$$

где $G_{\cup} = \langle V_{\cup}, X_{\cup} \rangle$, $V_{\cup} = \bigcup_{i=1}^N V_i$, $X_{\cup} = \bigcup_{i=1}^N X_i$.

При объединении помеченных вершин в объединенном тезаурусе возможно существование подмно-

жества дескрипторов, входящих более чем в один тезаурус. Кроме того, в связи с тем, что в различных тезаурусах одинаковые дескрипторы могут быть соединены связями разных типов, правило объединения дуг (в целях сохранения всех возможных типов связей тезаурусов) разрешает существование дуг, инцидентных одной и той же паре вершин, что приводит к формированию G_U как мультиграфа.

Для исследования свойств G_U введем следующие определения.

Назовем w_j -окрестностью вершины $v_k^i \in V_i$ подграф $G_i D_k^{w_j} \subset G_i$, включающий вершину v_k^i , все исходящие из нее дуги весом w_j и инцидентные им вершины (рис. 1).

Будем говорить, что вершины v_k^i и v_l^m «семантически похожи» по w_j -связям, если $v_k^i = v_l^m$ и $G_i D_k^{w_j} \cap G_m D_l^{w_j} \neq P$, где P – пустой или тривиальный граф (рис. 2).

На рисунке представлен пример w_j -окрестностей двух равных вершин. В результате объединения будет получен подграф, содержащий вершины v_1 , v_2 и инцидентную им дугу w_j .

3. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ОБЪЕДИНЕННОГО ТЕЗАУРУСА

Для создания базы данных экспериментального объединенного тезауруса использована концептуальная модель, представленная на рис. 3.

Сущность «Дескриптор» предназначена для хранения перечня уникальных дескрипторов со свойством, определяющим длину дескриптора в словах.

Сущность «Дескриптор тезауруса» позволяет в рамках модели сохранить данные о принадлежности отдельного дескриптора всем тезаурусам, которые его содержат. Свойство «максимальный уровень вложенности» определяет максимальное расстояние дескриптора от корневого в отдельном тезаурусе, а свойство «максимальная длина цепочки» – максимальную длину маршрута (от корня до листа), содержащего дескриптор в отдельном тезаурусе.

Отношение «Связь» предназначено для хранения типов связей. Для целей настоящей статьи были рассмотрены иерархические (ВТ – выше, НТ – ниже) и ассоциативные (РТ) связи.

Для формирования экспериментального объединенного тезауруса использовались тезаурусы разного назначения и разных областей знаний:

1. Тезаурусы по общественным наукам ИНИОН РАН – Лингвистика, Литература, Государство и право, Экономика, Религиоведение, Науковедение.
2. Тезаурусы РОСС (Русский общесемантический словарь) – Информатика, Финансовый, Общий, Локативный.
3. Тезаурус Международной системы ядерной информации ИНИС МАГАТЭ.

Построенный путем загрузки в реляционную СУБД исходных тезаурусов экспериментальный объединенный тезаурус включает 50538 уникальных лексических единиц (ЛЕ), в том числе 37771 дескрипторов и 12767 аскрипторов.

Тезаурусы ИНИОН РАН и ИНИС представлены двумя типами лексических единиц – дескрипторами и аскрипторами, тезаурусы РОСС содержат только дескрипторы.

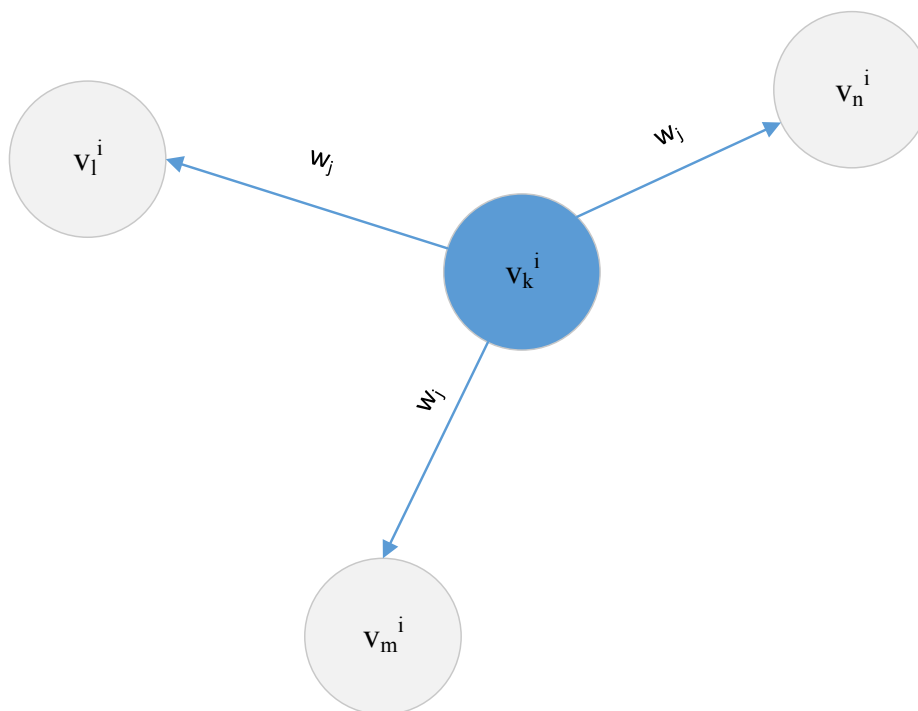


Рис. 1. w_j -окрестность вершины v_k^i

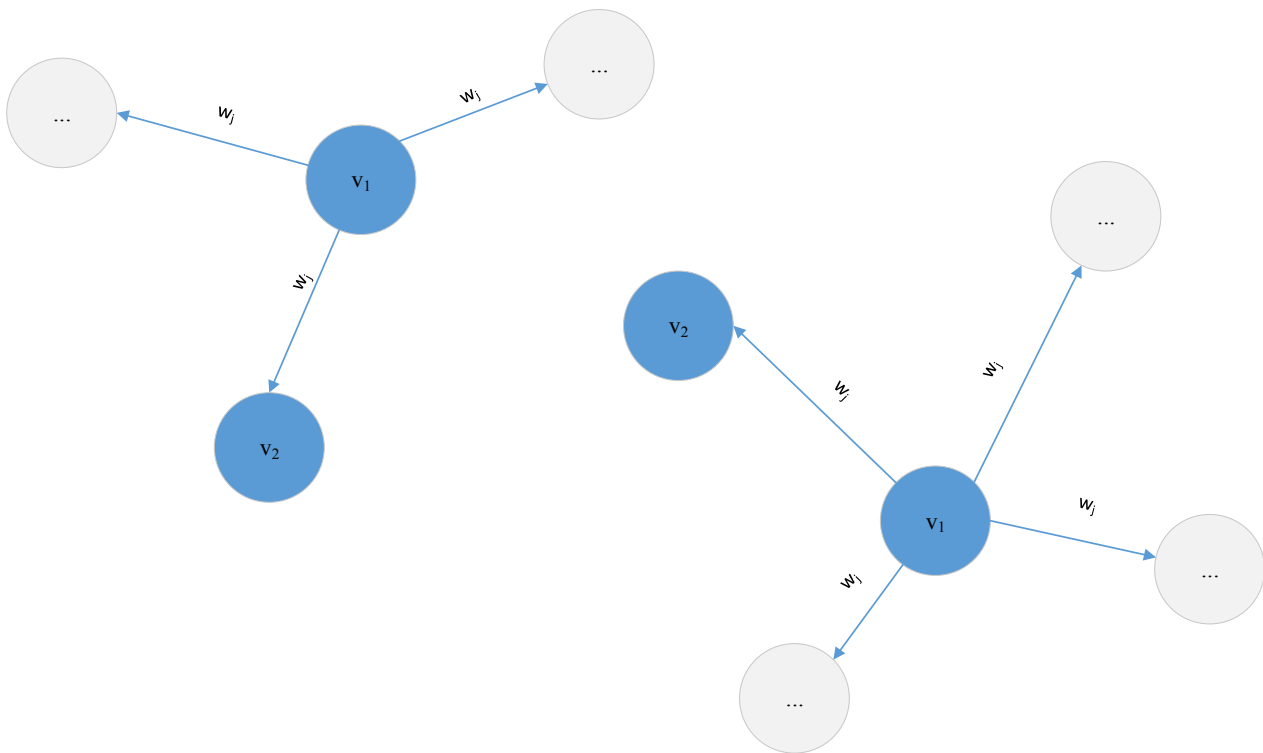


Рис. 2. Вершины, «семантически похожие» по W_j -связям

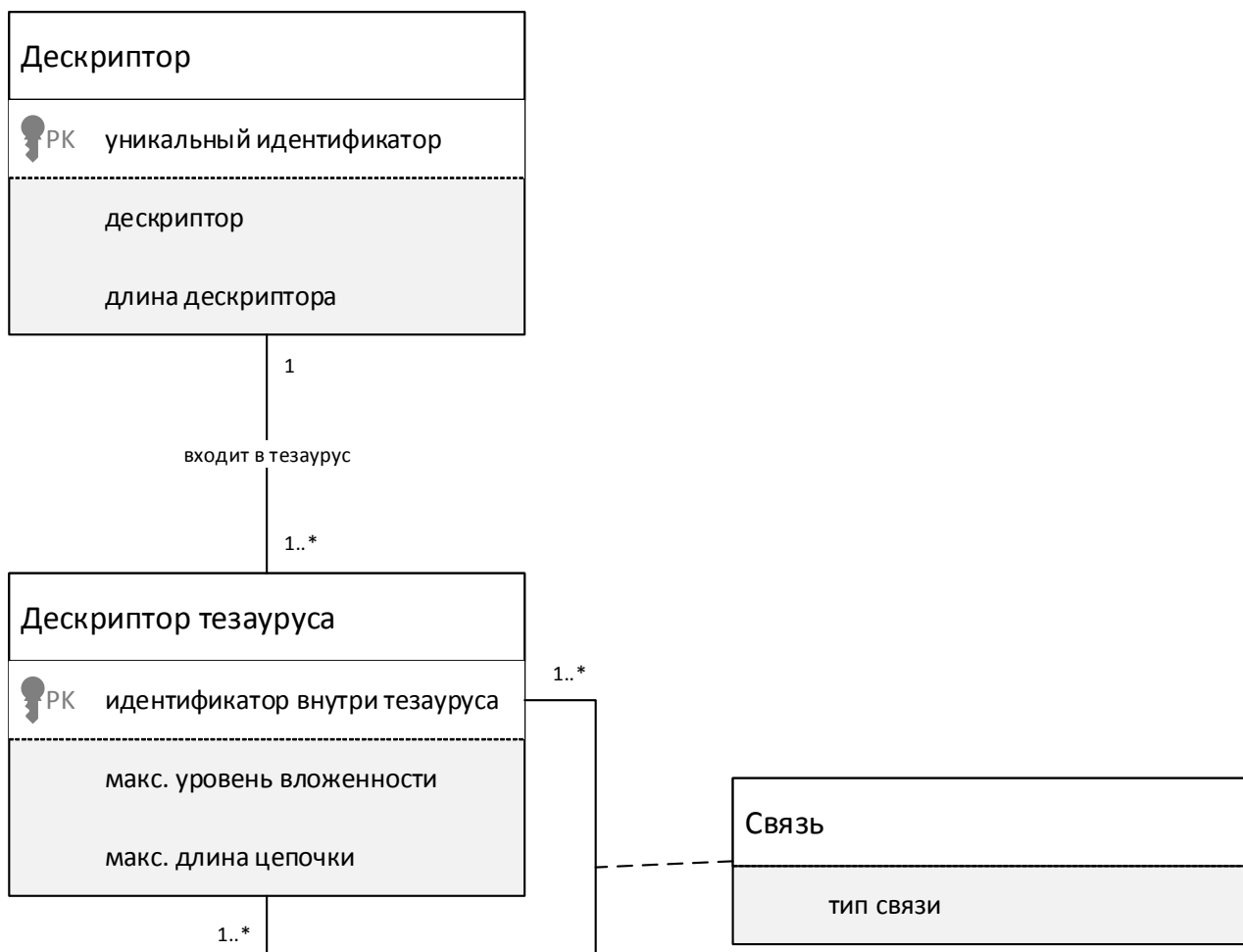


Рис. 3. ER-диаграмма объединенного тезауруса

В табл. 1 приведены основные количественные и структурные характеристики тезаурусов:

- общее количество ЛЕ;
- количество дескрипторов;
- количество аскрипторов;
- максимальное число уровней иерархических деревьев;
- количество дескрипторов верхнего уровня (корневых);
- количество дескрипторов нижнего уровня (листьев);
- количество дескрипторов без иерархических связей.

По структурным характеристикам среди анализируемых тезаурусов можно выделить тезаурусы с относительно большим количеством дескрипторов без видового развития (что, в частности, характеризуется тем, что около 25% и более дескрипторов не имеют иерархических связей, например, тезаурус «Экономика»). С другой стороны, выделяются тезаурусы, не имеющие дескрипторов без иерархических связей (например, тезаурусы РОСС) или содержащие менее 10% дескрипторов без иерархических связей (например, «Лингвистика» и «Литература»). Отметим также, что по глубине развития Тезаурусы РОСС имеют наименьшую детализацию понятий.

Количество дескрипторов верхнего уровня во всех рассматриваемых тезаурусах менее 10% от общего количества ЛЕ, количество дескрипторов-листьев в основном превышает 50% (менее 50% листьев наблюдается у тезаурусов, содержащих более 10% дескрипторов без видового развития).

Структурная неоднородность (дерева и лексических единиц) отражена распределениями, представленными на рис. 4 и рис.5.

Исследование структуры родо-видовых деревьев, соответствующих отдельным корневым дескрипторам тезаурусов, дало следующие результаты:

1. Тезаурусам свойственна существенная неравномерность глубины развития (детализации, представленной длиной цепочки дескрипторов) корневых понятий, что иллюстрирует рис. 4, на котором представлены графики распределения процента корневых дескрипторов по числу родо-видовых уровней. В подавляющем большинстве исследованных тезаурусов (за исключением тезаурусов «Общий» и «Финансовый») около 50% и более корневых дескрипторов порождают двухуровневые родо-видовые деревья (корень – листья).

2. Ранговое распределение корневых дескрипторов по мощности их родо-видовых деревьев (рис. 5 – распределение первых 20 корневых дескрипторов) показывает наличие трех групп тезаурусов:

- тезаурусы, в которых уже одно родо-видовое дерево содержит более 50% всех дескрипторов («Лингвистика», «Литература», «Финансовый», «Информатика»);
- тезаурусы с мощностью родо-видового дерева от 15% до 30% всех дескрипторов («Науковедение», «Религиоведение», «Общий», «Локативный», INIS);
- тезаурусы с относительно равномерным распределением дескрипторов по родо-видовым деревьям («Экономика», «Право»).

Таблица 1

Количественные и структурные характеристики тезаурусов

№ п/п	Тезаурус	Кол-во ЛЕ	Кол-во дескрипторов	Кол-во аскрипторов	Максимальное число родо-видовых уровней	Дескрипторы без иерархических связей		Дескрипторы верхнего уровня		Дескрипторы-листья	
						Кол-во	% от ЛЕ	Кол-во	% от ЛЕ	Кол-во	% от ЛЕ
1	Лингвистика	1887	1853	34	14	68	3,60	56	2,97	1301	68,95
2	Литература	734	734	0	11	61	8,31	42	5,72	488	66,48
3	Право	8079	3405	4674	8	871	10,78	337	4,17	1734	21,46
4	Экономика	3706	3407	299	9	1405	37,91	269	7,25	1472	39,72
5	Религиоведение	1238	1179	59	10	319	35,77	39	3,15	673	54,36
6	Науковедение	4053	3991	62	23	1004	24,77	377	9,30	1766	43,57
7	Информатика	663	663	0	12	0	0	56	8,45	340	51,28
8	Финансовый	2945	2945	0	10	0	0	16	0,54	2277	77,32
9	Общий	1642	1642	0	4	0	0	20	1,21	1590	96,89
10	Локативный	408	408	0	6	0	0	20	4,90	334	81,86
11	INIS	30418	21800	8618	12	2346	7,71	906	2,98	15889	52,24

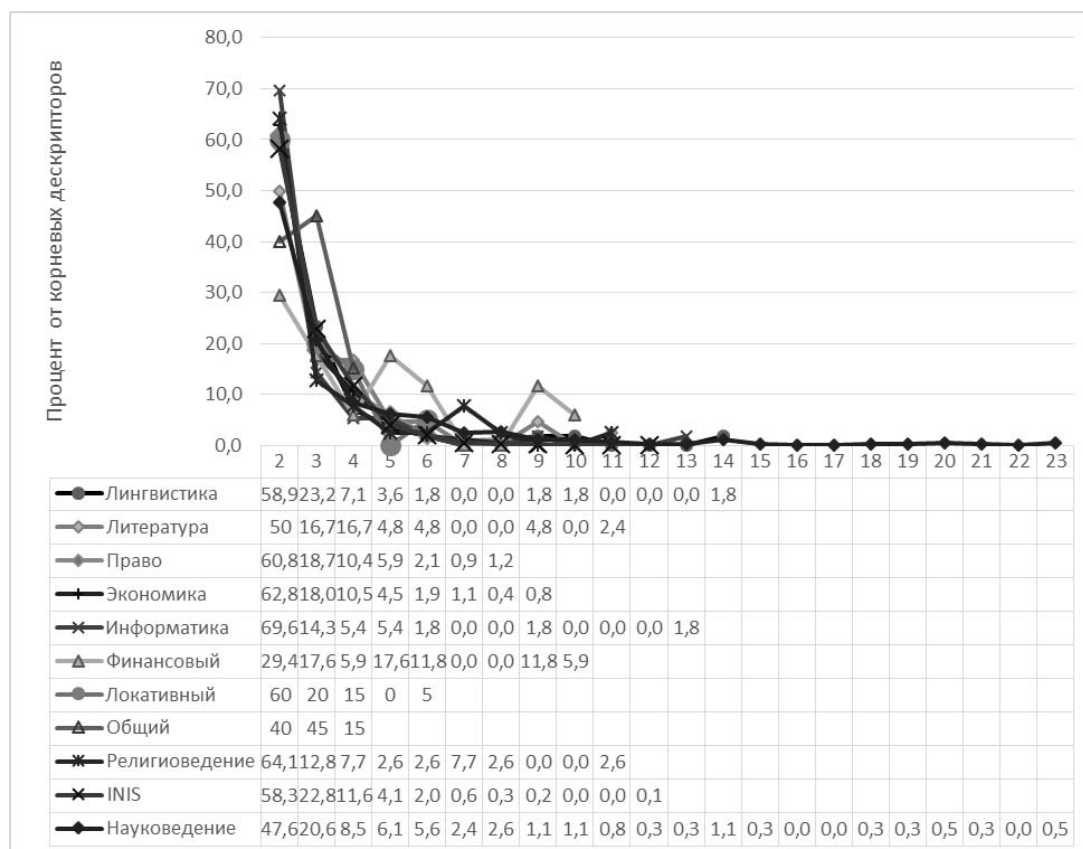


Рис. 4. Распределение корневых дескрипторов по уровням родо-видовых деревьев

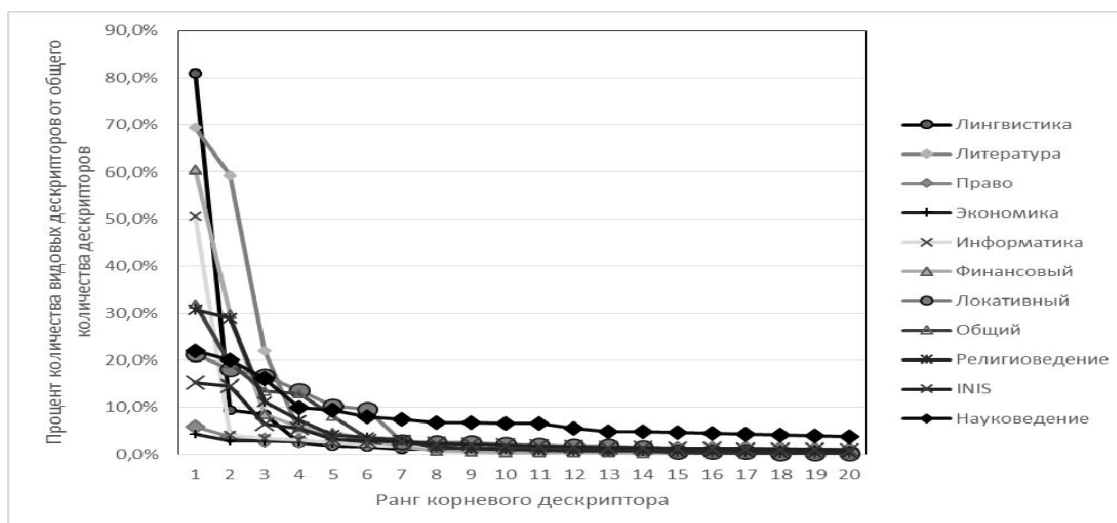


Рис. 5. Ранговое распределение корневых дескрипторов

Распределения лексических единиц по их длине (в словах) приведены в табл. 2.

Сумма одно- и двухсловных ЛЕ составляет в среднем 85,54% (минимум 69,40% в тезаурусе INIS, максимум 99,21% в Общем тезаурусе РОСС). Соотношения между количеством терминов разной длины для разных тезаурусов разные, но заметно зависят от характера предметной области: например, в тезаурусе INIS наблюдается преобладание дескрипторов длиной два и более слов.

4. АНАЛИЗ БЛИЗОСТИ ТЕЗАУРУСОВ

Для анализа близости тезаурусов будем рассматривать множество «политематических» дескрипторов, которые входят более чем в один тезаурус. В экспериментальном объединенном тезаурусе таких дескрипторов 3908 – 10,35% от общего количества. Распределение числа дескрипторов в зависимости от числа содержащих их тезаурусов приведено в табл. 3.

Распределение ЛЕ по длине (в словах)

№ п/п	Тезаурус	Макс. длина дескриптора	ЛЕ длиной 1		ЛЕ длиной 2		ЛЕ длиной 3		ЛЕ длиной 4		ЛЕ длиной > 4	
			Кол-во	%	Кол-во	%	Кол-во	%	Кол-во	%	Кол-во	%
1	Лингвистика	5	790	41,87	893	47,32	185	9,80	17	0,90	2	0,11
2	Литература	5	353	48,09	311	42,37	56	7,63	10	1,36	4	0,55
3	Право	10	1637	20,26	4208	52,09	1547	19,48	471	5,83	216	2,67
4	Экономика	7	898	24,23	2208	59,58	464	12,52	99	2,67	37	1,00
5	Религиоведение	4	834	67,37	348	28,11	49	3,96	7	0,56	0	0
6	Науковедение	5	1516	37,40	2188	53,98	312	7,70	32	0,79	5	0,12
7	Информатика	6	168	25,34	347	52,34	109	16,44	29	4,37	10	1,51
8	Финансовый	8	299	10,15	1850	62,82	548	18,61	186	6,31	62	2,11
9	Общий	3	1565	95,31	64	3,90	13	0,79	0	0	0	0
10	Локативный	3	317	77,70	87	21,32	4	0,98	0	0	0	0
11	INIS	11	5897	19,39	15213	50,01	4807	15,80	2578	8,47	1923	6,32

Таблица 3

Распределение числа дескрипторов в зависимости от числа содержащих их тезаурусов

Число тезаурусов	7	6	5	4	3	2	1
Число общих дескрипторов	2	14	50	190	731	2921	33863

Приведенное в табл. 4 распределение по длине дескрипторов объединенного тезауруса и дескрипторов, входящих только в один тезаурус, показывает существенное преобладание словосочетаний из двух слов, а в сумме дескрипторы длиной в одно и два слова составляют около 80% от общего числа.

Около 90% «политематических» дескрипторов (т.е. входящих более чем в один тезаурус) имеют длину не более двух слов, при этом количество однословных и двухсловных дескрипторов уже практически совпадает, распределение терминов по длине в подмножестве политематических дескрипторов сместилось к меньшим значениям.

Для количественного исследования близости тезаурусов на основе анализа состава дескрипторов и связей между ними введем следующие коэффициенты.

$$S_{ij}^v = \frac{|V_i \cap V_j|}{|V_i \cup V_j|}, \quad (1)$$

$$S_{ij}^x = \frac{|X_i \cap X_j|}{|X_i \cup X_j|}, \quad (2)$$

S_{ij}^v – коэффициент лексической близости; показывает долю общих дескрипторов в попарном пересечении тезаурусов.

S_{ij}^x – коэффициент семантической близости на уровне контекста; характеризует долю общих связей в попарном пересечении тезаурусов, что отражает «осмысление» термина контекстом за счет связанных дескрипторов.

Значения коэффициентов (1) и (2), рассчитанные для исследуемых тезаурусов, приведены в табл. 5. В верхней треугольной части таблицы приведены значения S_{ij}^v , в нижней треугольной части – значения S_{ij}^x .

Проведенные расчеты показывают, что для любой пары тезаурусов $S_{ij}^x \leq S_{ij}^v$, т.е. значение коэффициента семантической близости всегда меньше значения коэффициента лексической близости (за исключением случаев, когда $S_{ij}^x = S_{ij}^v = 0$). Это подтверждает вывод, сделанный еще в [14]: «...одно только совпадение терминов еще не дает решения проблемы совместимости словарей, если одновременно не нормализована система понятий, представленная парадигматическими отношениями». То есть, в данном случае объединение тезаурусов по правилам объединения графов будет недостаточно – необходимо иметь дополнительно средства выявления и учета полисемии.

Таким образом, необходимо различать два типа полисемии дескриптора: *парадигматическую полисемию*, т.е. многозначность на уровне соответствующих термину понятий, и *синтагматическую полисемию* – многоаспектность его использования. Полисемия первого типа при выполнении операции объединения тезаурусов должна приводить к дублированию дескриптора с обязательным использованием реляторов или лексических примечаний [3]. В случае полисемии второго типа необходимо иметь возможность указания при связях дескриптора аспекта его использования и требует введения дополнительных атрибутов для связей, что, в частности, не предусмотрено в [3].

В табл. 6 представлено количество ненулевых значений S_{ij}^v и S_{ij}^x для каждого из тезаурусов.

Как видно из табл. 6, в большинстве случаев в тезаурусах отмечается наличие одних и тех же дескрипторов с различным понятийным контекстом (т.е. с непересекающимися или частично пересекающимися w_j -окрестностями) или наличие разных типов связей между одинаковыми дескрипто-

рами. Напомним, именно наличие в разных тезаурусах для одной и той же пары дескрипторов связей разного типа является признаком синтагматической полисемии. Для рассматриваемых тезаурусов было рассчитано процентное соотношение таких случаев к общему количеству связей между одинаковыми дескрипторами (табл. 7).

Таблица 4

Распределение дескрипторов по длине (в словах)

Тезаурус	Макс. длина дескриптора	Дескрипторы длиной 1		Дескрипторы длиной 2		Дескрипторы длиной 3		Дескрипторы длиной 4		Дескрипторы длиной > 4	
		Кол-во	%	Кол-во	%	Кол-во	%	Кол-во	%	Кол-во	%
Объединенный тезаурус	11	9760	25,84	20132	53,30	4860	12,87	2065	5,47	954	2,52
Дескрипторы более чем 1 тезауруса	6	1720	44,01	1734	44,37	353	9,03	80	2,05	21	0,54
Дескрипторы 1 тезауруса	11	8040	23,74	18398	54,33	4507	13,31	1985	5,86	933	2,75

Таблица 5

Коэффициенты лексической и семантической близости

	Лингвистика	Литература	Право	Экономика	Религиоведение	Науковедение	Информатика	Финансовый	Общий	Локативный	INIS
Лингвистика		0,050	0,006	0,012	0,017	0,031	0,004	0,001	0,004	0,000	0,001
Литература	0,008		0,006	0,016	0,030	0,028	0,001	0,001	0,003	0,000	0,000
Право	0,000	0,000		0,190	0,014	0,074	0,001	0,027	0,011	0,002	0,009
Экономика	0,001	0,001	0,183		0,032	0,110	0,001	0,014	0,008	0,002	0,006
Религиоведение	0,000	0,001	0,001	0,003		0,038	0,000	0,001	0,005	0,002	0,001
Науковедение	0,002	0,002	0,011	0,013	0,004		0,004	0,010	0,025	0,001	0,013
Информатика	0,000	0,000	0,000	0,000	0,000	0,000		0,000	0,003	0,002	0,000
Финансовый	0,000	0,000	0,002	0,001	0,000	0,001	0,000		0,004	0,000	0,001
Общий	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001		0,003	0,005
Локативный	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000		0,003
INIS	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,000	0,000	

Таблица 6

Количественное распределение ненулевых значений коэффициентов

	Лингвистика	Литература	Право	Экономика	Религиоведение	Науковедение	Информатика	Финансовый	Общий	Локативный	INIS
$S_{ij}^y > 0$	9	8	10	10	9	10	7	8	10	7	9
$S_{ij}^x > 0$	3	4	5	6	4	7	0	3	1	0	2

Процентное соотношение связей разного типа к общему количеству связей между одинаковыми дескрипторами

	Лингвистика	Литература	Право	Экономика	Религиоведение	Науковедение	Информатика	Финансовый	Общий	Локативный	INIS
Лингвистика											
Литература	22,7										
Право	33,3	42,9									
Экономика	27,3	22,2	6,5								
Религиоведение	50,0	45,5	30,0	26,3							
Науковедение	30,4	18,2	16,9	20,1	22,2						
Информатика	0	0	0	0	0	0					
Финансовый	0	0	13,2	38,1	0	9,7	0				
Общий	0	0	0	0	0	0	0	0			
Локативный	0	0	0	0	0	0	0	0	0		
INIS	100,0	0	2,0	13,6	0	26,6	0	50,0	66,6	0	

5. ИССЛЕДОВАНИЕ МНОЖЕСТВА ОБЩИХ ДЕСКРИПТОРОВ

Во множестве дескрипторов, входящих более чем в один тезаурус, было выделено подмножество дескрипторов, которые «семантически похожи» с точки зрения теоретико-графового представления, т.е. имеют одинаковые связи хотя бы по одному типу (BT, NT или RT) во всех тезаурусах. В табл. 8 приведены распределения подмножеств дескрипторов по типам связей с разбивкой по их длине (в словах) относительно исходного подмножества «Всего семантически похожих».

Как видно из таблицы, «семантически похожи» по какому-либо типу связи около 40% всех дескрипторов, однако среди однословных дескрипторов их чуть более 20%, а среди дескрипторов длиной три и более слова – таких уже 80%. Данные по отдельным типам связи также демонстрируют возрастание доли «семантически похожих» дескрипторов с увеличением длины дескриптора. Следует отметить и существенное преобладание похожих ассоциативных связей над иерархическими.

Для характеристики дескрипторов с точки зрения их «координатного» положения в тезаурусах рассмотрены следующие непересекающиеся подмножества:

S_1 – дескрипторы, не имеющие иерархических связей ни в одном из тезаурусов; $|S_1|=550$;

S_2 – дескрипторы, являющиеся корневыми во всех тезаурусах; $|S_2|=105$;

S_3 – дескрипторы, являющиеся листьями иерархических деревьев во всех тезаурусах; $|S_3|=907$;

S_4 – дескрипторы, имеющие во всех тезаурусах высстоящие и нижестоящие связи, т.е. те дескрипторы, которые находятся на промежуточных иерархических уровнях; $|S_4|=150$.

S_5 – дескрипторы, не вошедшие в подмножества $S_1 - S_4$, т.е. занимающие разное положение в структурах разных тезаурусов; $|S_5|=2196$.

Максимальное количество дескрипторов вошли в подмножество S_5 . Это означает, что одна и та же лексическая единица в разных тезаурусах имеет разное контекстное окружение. В подмножества $S_1 - S_4$ входят дескрипторы (менее половины всех исследуемых), занимающие одинаковое положение во всех содержащих их тезаурусах.

Такое разнообразие понятийного состояния лексической единицы в различных тематических областях может приводить к следующим практическим выводам и рекомендациям:

1. Объем понятия может быть удачно расширен путем объединения всех связей тезаурусов.

2. Необходимо зафиксировать и разрешить на уровне связей лексической единицы ситуацию синтагматической полисемии.

3. Необходимо зафиксировать и разрешить на уровне самой лексической единицы ситуацию парадигматической полисемии.

Для анализа полисемии в каждом из подмножеств $S_1 - S_4$ были выявлены дескрипторы, «семантически похожие» по определенным связям во всех тезаурусах – т.е. дескрипторы, объем понятия которых предположительно может быть расширен.

В табл. 9 приведены количественные распределения дескрипторов по длине (в словах) и рассчитана (в процентах) доля дескрипторов, вошедших в подмножества «семантически похожих», в общем числе дескрипторов соответствующего подмножества.

Более 40% «политематических» дескрипторов в каждом из подмножеств «семантически похожи» по какому-либо типу связи, что выше уровня общих показателей. При этом «семантически похожих» по иерархическим связям более 57% (максимальное количество – среди корневых дескрипторов). Заметим также, что возрастание доли «семантически похожих» дескрипторов с увеличением длины дескриптора сохраняется, т.е. выводы о необходимости учета полисемии применимы, прежде всего, к однословным дескрипторам.

Распределение «семантически похожих» дескрипторов по длине (в словах)

Подмножество дескрипторов	Кол-во	Длина дескриптора			
		1	2	3	>3
«Всего семантически похожих»	3908	1720	1734	353	101
«похожие» по ВТ-, NT- или RT-связям	1548	370	868	232	81
% от «всего»	39,61	21,51	50,06	65,72	80,20
«похожие» по ВТ-связям	767	103	475	139	50
% от «всего»	19,63	5,99	27,39	39,38	49,50
«похожие» по NT-связям	224	66	100	37	21
% от «всего»	5,73	3,84	5,77	10,48	20,79
«похожие» по RT-связям	1286	290	712	209	75
% от «всего»	32,91	16,86	41,06	59,21	74,26

Таблица 9

Доля «семантически похожих» дескрипторов

Под-множе-ство	Дескрипторы	Всего	Длина дескриптора			
			1	2	3	>3
S_1	Всего в подмножестве	550	223	241	62	24
	«похожие» по RT-связям	236	69	117	35	15
	% от «всего»	42,9	30,9	48,6	56,5	62,5
S_2	Всего в подмножестве	105	61	32	10	2
	«похожие» по NT-связям	70	35	25	8	2
	% от «всего»	66,7	57,4	78,1	80,0	100,0
S_3	Всего в подмножестве	907	309	442	121	35
	«похожие» по ВТ-связям	523	66	320	105	32
	% от «всего»	57,7	21,4	72,4	86,8	91,4
S_4	Всего в подмножестве	150	30	75	30	15
	«похожие» по ВТ-связям	109	11	58	25	15
	«похожие» по NT-связям	99	10	52	24	13
	«похожие» по ВТ- и NT-связям	86	6	45	22	13
	% от «всего»	57,3	50,0	86,7	90,0	88,7

Подмножество S_5 составили дескрипторы, занимающие в тезаурусах разное положение. Всего в подмножестве S_5 – 2196 дескрипторов, причем около 93% из них имеют длину одно или два слова.

Для выявления «семантически похожих» дескрипторов подмножество S_5 было представлено (рис. 6) как объединение трех множеств – H , L и N , где

H – множество дескрипторов, являющихся корневыми хотя бы в одном из тезаурусов; $|H| = 585$;

L – множество дескрипторов, которые являются листьями хотя бы в одном из тезаурусов; $|L| = 1504$;

N – множество дескрипторов, у которых хотя бы в одном из тезаурусов отсутствуют иерархические связи; $|N| = 1540$.

Объединение представлено следующими непересекающимися подмножествами – структурными группами:

$M_1 = H \cap L \cap N$ – дескрипторы, занимающие в разных тезаурусах положение корневых, листьев и дескрипторов без иерархических связей;

$M_2 = (H \cap L) \setminus M_1$ – дескрипторы, являющиеся корневыми в одних тезаурусах и листьями – в других;

$M_3 = (N \cap L) \setminus M_1$ – дескрипторы, являющиеся либо листьями, либо дескрипторами без иерархических связей;

$M_4 = (H \cap N) \setminus M_1$ – дескрипторы, являющиеся либо корневыми, либо дескрипторами без иерархических связей;

$M_5 = H \setminus (L \cup N)$ – дескрипторы, занимающие в разных тезаурусах положение корневых или промежуточных в иерархических деревьях;

$M_6 = N \setminus (L \cup H)$ – дескрипторы либо промежуточные в иерархических деревьях, либо не имеющие иерархических связей;

$M_7 = L \setminus (H \cup N)$ – дескрипторы, занимающие в разных тезаурусах положение листьев или промежуточных в иерархических деревьях.

В табл. 10 приведены мощности подмножеств и количественные распределения дескрипторов по длине.

Доля «семантически похожих» дескрипторов в подмножествах $M_1 - M_7$, выделенных на S_5 , существенно ниже, чем в подмножествах $S_1 - S_4$. Следует отметить, что дескрипторы «семантически похожие» только по РТ-связям, наблюдаются во всех подмножествах $M_1 - M_7$. Дескрипторы, «семантически похожие» по НТ-связям, обнаружены только в подмножестве M_5 , а «семантически похожие» по ВТ-связям – только в M_7 . В табл. 11 приведены результаты распределения дескрипторов, «семантически похожих» по РТ-связям, в подмножествах $M_1 - M_7$.

В каждой из структурных групп таких дескрипторов менее 20%, причем лидируют дескрипторы длиной два и три слова, а однословных дескрипторов – менее 15%.

В табл. 12 представлена характеристика дескрипторов, «семантически похожих» по НТ-связям (подмножество M_5) и ВТ-связям (подмножество M_7).

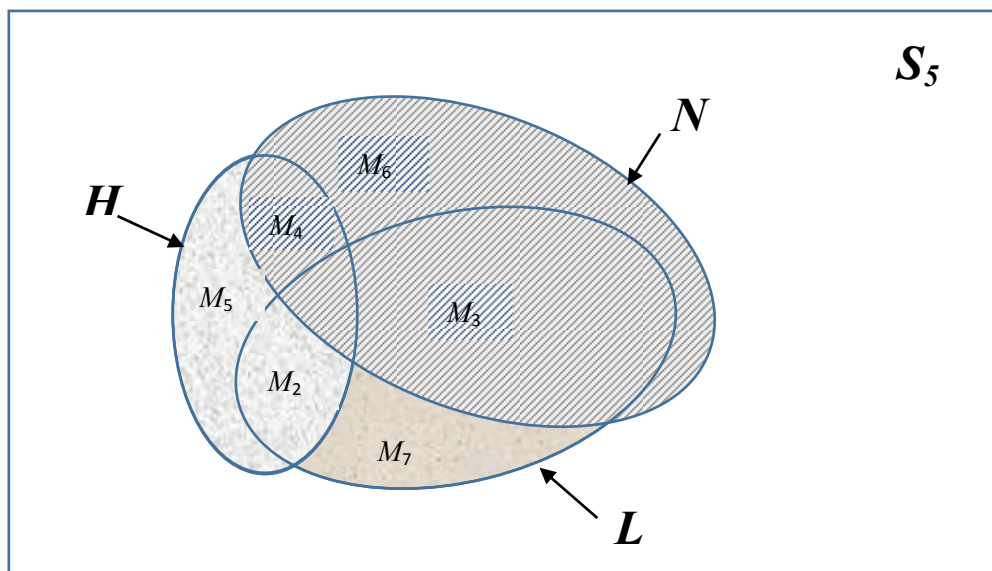


Рис. 6. Разбиение подмножества S_5

Таблица 10

Распределение по длине дескрипторов подмножеств

Подмножество	Кол-во дескрипторов	Дескрипторы длиной 1 слово		Дескрипторы длиной 2 слова		Дескрипторы длиной 3 слова		Дескрипторы длиной >3 слов	
		Кол-во	%	Кол-во	%	Кол-во	%	Кол-во	%
M_1	62	44	70,97	14	22,58	3	4,84	1	1,61
M_2	138	84	60,87	46	33,33	8	5,80	0	0
M_3	920	418	45,43	434	47,17	59	6,41	9	0,98
M_4	251	161	64,14	74	29,48	12	4,78	4	1,59
M_5	134	76	56,72	45	33,58	7	5,22	6	4,48
M_6	307	152	49,51	131	42,67	22	7,17	2	0,65
M_7	384	162	42,19	200	52,08	19	4,95	3	0,78
Всего	2196	1097	49,95	944	42,99	130	5,92	25	1,14

Распределение дескрипторов, «семантически похожих» по RT-связям

Подмножество	Дескрипторы	Всего	Длина дескриптора			
			1	2	3	>3
M_1	Всего	62	44	14	3	1
	«похожие» по RT-связям	5	1	3	1	0
	% от «всего»	8,06	2,27	21,43	33,33	0
M_2	Всего	138	84	46	8	0
	«похожие» по RT-связям	27	11	11	5	0
	% от «всего»	19,57	13,10	23,91	63,5	0
M_3	Всего	920	418	434	59	9
	«похожие» по RT-связям	172	54	93	22	3
	% от «всего»	18,70	12,92	21,43	37,29	33,33
M_4	Всего	251	161	74	12	4
	«похожие» по RT-связям	44	15	24	6	2
	% от «всего»	17,53	9,32	32,43	50,00	50,00
M_5	Всего	134	76	45	7	6
	«похожие» по RT-связям	13	7	5	1	0
	% от «всего»	9,70	9,21	11,11	14,29	0
M_6	Всего	307	152	131	22	2
	«похожие» по RT-связям	49	17	24	6	2
	% от «всего»	15,96	11,18	18,32	27,27	100
M_7	Всего	384	162	200	19	3
	«похожие» по RT-связям	31	10	20	1	0
	% от «всего»	8,07	6,17	20,00	5,26	0

Таблица 12

Распределение дескрипторов, «семантически похожих» по NT- и VT-связям

Подмножество	Дескрипторы	Всего	Дескрипторов длиной			
			1	2	3	>3
M_5	Всего	134	76	45	7	6
	«похожие» по NT-связям	55	21	23	5	6
	«похожие» по NT- и RT- связям	38	8	19	5	6
	% от «всего»	41,04	27,63	51,11	71,43	100
M_7	Всего	384	162	200	19	3
	«похожие» по VT-связям	135	26	97	9	3
	«похожие» по VT- и RT- связям	51	9	34	5	3
	% от «всего»	35,16	16,05	48,50	47,37	100

Проведенный анализ подтверждает, что и структура тезауруса, и правила описания лексических единиц и связей играют существенную роль в повышении качества решения задач информационного поиска [15]. При этом наиболее чувствительны к семантике контекста лексические единицы длиной в одно слово.

Проблемы, вызываемые парадигматической или синтагматической полисемией лексических единиц, могут быть решены путем добавления атрибута тематической характеристики к свойствам как дескриптора, так и конкретной связи. Дескриптор как лексическая

единица при этом «размножается» с добавлением релятора, значение которого может соответствовать, например, номеру главного класса фасета «Общий предмет» УДК (на международном уровне) или коду верхнего уровня ГРНТИ (на национальном уровне). Отметим, что предложения [16] по структурным параметрам тезауруса (на примере расширения запросов к машинам поиска Интернета) предопределяют для лексической единицы характеристику, роль которой аналогична роли релятора. Приведенные в настоящем исследовании данные показывают необходимость ре-

лятора и для связи, причем при формировании политематических лексических ресурсов целесообразно использование таких объектов лингвистического обеспечения, как библиотечно-библиографические классификации и тематические рубрикаторы.

ЗАКЛЮЧЕНИЕ

Для сравнительного анализа понятийно-лексического состава и семантических связей специализированных тезаурусов была построена база данных экспериментального объединенного тезауруса, для чего были использованы тезаурусы разного назначения (информационно-поисковые и лингвистические) и относящиеся к разным областям знаний (гуманитарные и технические).

По структурным показателям (отражающим, в частности, систематизированность и глубину проработанности предметной области) среди анализируемых тезаурусов выделяются тезаурусы с относительно большим количеством дескрипторов без видового развития (25% и более дескрипторов не имеют иерархических связей, например, тезаурус «Экономика»), а с другой стороны – тезаурусы не имеющие дескрипторов без иерархических связей (например, тезаурусы РОСС, имеющие также заметно меньшую детализацию понятий) или содержащие менее 10% дескрипторов без иерархических связей (например, «Лингвистика» и «Литература»). В тезаурусах преобладают одно- и двухсловные дескрипторы, хотя соотношения между этими группами в разных тезаурусах различно.

Из 37771 уникальных дескрипторов, составляющих базу экспериментального объединенного тезауруса, 10,35% являются «политематическими», т.е. входят более чем в один тезаурус. Около 90% из них имеют длину не более двух слов, при этом количество однословных и двухсловных дескрипторов приблизительно одинаково.

Около 40% всех «политематических» дескрипторов «семантически похожи» по какому-либо типу связи, при этом наблюдается возрастание доли «семантически похожих» дескрипторов с увеличением длины дескриптора.

Выявлено существенное преобладание числа похожих ассоциативных связей над иерархическими, причем «похожесть» только для этого типа связи не зависит от иерархического положения дескриптора (корневой, промежуточный, лист).

Исходя из обобщенной графовой модели тезауруса предложено различать два типа полисемии: парадигматическую полисемию, т.е. многозначность на уровне понятий, соответствующих термину, и синтагматическую полисемию – многоаспектность его использования. При полисемии первого типа в объединенном тезаурусе необходимо дублировать дескриптор с обязательным использованием реляторов или лексических примечаний, а в случае синтагматической полисемии при связях дескриптора необходимо указывать аспект. Некоторые примеры парадигматической и синтагматической полисемии для структурных подмножеств приведены в Приложении.

СПИСОК ЛИТЕРАТУРЫ

1. Михайлов А.М., Черный А.И., Гиляревский Р.С. Основы информатики. – М.: Наука, 1968. – 756 с.
2. Арапов М.В. Некоторые принципы построения словаря типа «тезаурус» // Научно-техническая информация. – 1964. – № 4. – С. 40.
3. ГОСТ 7.25 – 2001. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления. – М.: Изд-во стандартов, 2001.
4. Информационно-поисковый тезаурус по информатике / сост. Н.А. Пашенко, Е.Б. Ксенофонтова, В.Ф. Скоробогатая. – М.: ВИНТИ, 1987. – 466 с.
5. Cheong H, Chiu I., Shu L.H., Stone R.B., Mcadams D.A. Biologically meaningful keywords for functional terms of the functional basis // Journal of Mechanical Design. Transactions of the ASME. – 2011. – Т. 133, № 2.
6. Жижимов О.Л., Скачков Д.М. О географическом поиске информации в «негеографических» информационных системах: использование ретроспективного тезауруса // Труды XIX Междунар. конф. «КРЫМ 2012. Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса» (Судак, 04–07 июня 2012 г.). – М.: ГПНТБ, 2012.
7. Suzuki Y., Fukumoto F. Multi-labeled patent document classification using technical term thesaurus // Proceedings of the International Joint Conference on Knowledge Engineering and Ontology Development – KEOD 2011. – Paris, 2011. – P. 105.
8. Thabit Sabbah, Ali Selamat, Mahmood Ashraf, Tutut Herawan. Effect of thesaurus size on schema matching quality // Knowledge-Based Systems. – 2014. – Vol. 71. – P. 211–226.
9. Xiangming Mu, Kun Lu, Hohyon Ryu. Explicitly integrating MeSH thesaurus help into health information retrieval systems: An empirical user study // Information Processing and Management. – 2014. – Vol. 50. – P. 24–40.
10. Lu Zhiong, Kim Won, Wilbur W. John. Evaluation of query expansion using MeSH in PubMed // Information Retrieval. – 2009. – Vol. 12(1). – P. 69–80.
11. Sondhi Parikkshit, Sun Jimeng, Zhai ChengXiang, Sorrentino Robert, Kohn Martin S. Leveraging medical thesauri and physician feedback for improving medical literature retrieval for case queries // Journal of the American Medical Informatics Association. – 2012. – Vol. 19. – P. 851–858.
12. Chakrabarty Sugato, Chougule Rahul, Lesperance Ronald M. Ontology-guided knowledge retrieval in an automobile assembly environment. – Int J Adv Manuf Technol, 44, 2009, pp. 1237–1249.
13. Borst T. Usage and impact of controlled vocabularies in a subject repository for indexing and retrieval // LIBER Quarterly. – 2012. – Vol. 21(3–4). – P. 445–453.

14. Кулик А.И. Информационные сети и языковая совместимость дескрипторных информационно-поисковых систем. – М.: Советское радио, 1977. – 248 с.
15. Агеев М.С., Добров Б.В., Лукашевич Н.В., Штернов С.В. УИС РОССИЯ в РОМИП-2008: поиск и классификация нормативных документов / Российский семинар по оценке методов информационного поиска. Дубна, 9 октября 2008 г. Труды РОМИП 2007–2008. – СПб: НУ ЦСИ, 2008.
16. Браславский П.И. Тезаурус для расширения запросов к машинам поиска Интернета: структура и функции // Компьютерная лингвистика и интеллектуальные технологии: III Междунар. конф. Диалог'2003 (Протвино, 11–16 июня 2003 г.). – М.: Наука, 2003. – С. 95–100.

ПРИЛОЖЕНИЕ

Примеры парадигматической полисемии

Подмножество S_1		
Дескриптор:	Автаркия	
Тезаурус:	«Религиоведение»	
Связь Дескрипторы:	RT	Экзегетика; священное писание
Тезаурус:	«Экономика»	
Связь Дескрипторы:	RT	Экономическая политика; регионализм; национальная экономика; изоляционизм
Подмножество S_2		
Дескриптор:	Споры	
Тезаурус:	«INIS»	
Связь Дескрипторы:	NT	Конидии; микроспоры; бактериальные споры; грибы
Тезаурус:	«Государство и право»	
Связь Дескрипторы:	NT	Экономические споры; коллективные трудовые споры; хозяйственные споры; трудовые споры
Подмножество S_3		
Дескриптор:	Окно	
Тезаурус:	«Информатика»	
Связь Дескрипторы:	BT	Объект экрана
Тезаурус:	«Общий»	
Связь Дескрипторы:	BT	Мебель
Подмножество S_4		
Дескриптор:	Лицо	
Тезаурус:	«INIS»	
Связь Дескрипторы:	BT	Голова
Связь Дескрипторы:	NT	Глаза; нос
Тезаурус:	«Лингвистика»	
Связь Дескрипторы:	BT	Грамматические категории
Связь Дескрипторы:	NT	Глагол; первое лицо; второе лицо; третье лицо
Подмножество M_1		
Дескриптор:	Управление	
Тезаурус:	«Лингвистика»	
Связь Дескрипторы:	BT	Подчинение
Связь Дескрипторы:	NT	Слабое управление; сильное управление; предложное управление; падежное управление; косвенное управление; беспредложное управление
Связь Дескрипторы:	RT	Валентность
Тезаурус:	«Экономика»	
Связь Дескрипторы:	RT	Централизация управления; централизм; участие в управлении; управленческая деятельность; социология управления; коллегиальность; революция менеджеров; администрация; органы управления; наука управления; право и управление; единоначалие; децентрализация управления; бюрократия; аппарат управления
Связь Дескрипторы:	NT	Политический менеджмент; государственное управление; кризисное управление; общественное управление; региональное управление; самоуправление; управление народным хозяйством; управление наукой; управление хаосом; местное управление

Примеры синтагматической полисемии

Подмножество S_1		
Дескриптор:	Будущее	
Тезаурус:	«Религиоведение»	
Связь Дескрипторы:	RT	Время; пророчество; божественное провидение; прошлое; теология надежды; эсхатология; предопределение; настоящее
Тезаурус:	«Экономика»	
Связь Дескрипторы:	RT	Политическая мифология; политическое воображение; политическое прогнозирование; прогнозирование; утопия; футурология
Подмножество S_2		
Дескриптор:	Выбор	
Тезаурус:	«Экономика»	
Связь Дескрипторы:	NT	Политический выбор
Тезаурус:	«Науковедение»	
Связь Дескрипторы:	NT	Выбор технологии; выбор научных проблем; выбор проектов; выбор теории; выбор профессии
Подмножество S_3		
Дескриптор:	Янтарь	
Тезаурус:	«INIS»	
Связь Дескрипторы:	BT	Прочие органические соединения
Тезаурус:	«Общий»	
Связь Дескрипторы:	BT	Драгоценный камень
Подмножество S_4		
Дескриптор:	Реализм	
Тезаурус:	«Науковедение»	
Связь Дескрипторы:	BT	Онтология
Связь Дескрипторы:	NT	Научный реализм
Тезаурус:	«Литература»	
Связь Дескрипторы:	BT	Литературные направления; художественный метод
Связь Дескрипторы:	NT	Социалистический реализм; критический реализм; магический реализм; неореализм; просветительский реализм; реалистические элементы
Подмножество M_1		
Дескриптор:	Аксиология	
Тезаурус:	«Лингвистика»	
Связь Дескрипторы:	BT	Философия языка
	«Науковедение»	
Связь Дескрипторы:	NT	Ценностный подход; ценностные критерии; ценности; общечеловеческие ценности; аксиологический подход
Связь Дескрипторы:	RT	Ценностная ориентация
Тезаурус:	«Экономика»	
Связь Дескрипторы:	RT	Базовые ценности; культурные ценности; консервативные ценности; исламские ценности; западные ценности; либеральные ценности; демократические ценности; политические ценности; азиатские ценности; духовные ценности; моральные ценности; национальные ценности; общие ценности; религиозные ценности; система ценностей; социальные ценности; традиционные ценности; христианские ценности; ценностная ориентация

Материал поступил в редакцию 27.02.15.

Сведения об авторах

ГОЛИЦЫНА Ольга Леонидовна – кандидат технических наук, доцент кафедры Системного анализа Национального исследовательского ядерного университета «МИФИ»
e-mail: OLGolitsina@YANDEX.RU

МАКСИМОВ Николай Вениаминович – доктор технических наук, профессор кафедры Системного анализа Национального исследовательского ядерного университета «МИФИ»
e-mail: NV-MAKS@YANDEX.RU

Е Цисун, С.Д. Шелов

О классификации номенов и номенклатурных наименований (на материале наименований товаров)

Анализируются номены и номенклатурные наименования на материале официальных наименований товаров, содержащихся в «Товарном словаре» 1956–1961 гг. издания. Выделяются различные типы номенов и номенклатурных наименований, и формулируются актуальные вопросы орфографии и пунктуации соответствующих обозначений.

Ключевые слова: номены, номенклатура, номенклатурные наименования, терминология, наименования товаров

Российскому терминоведению принадлежит заслуга наиболее разностороннего и глубокого обсуждения вопроса о разграничении терминологии и номенклатуры. Возвращаясь к европейской традиции обсуждения философских проблем языка науки в лице У. Уэвеля (1794–1866, в других написаниях Уэвелла, Хьюэлла) [1] и Дж. Ст. Милля (1806–1873) [2], первоначально Г.О. Винокур [3] и А.А. Реформатский [4, 5], а затем и другие авторы неоднократно обращаются к обсуждению различия (и противопоставления) номенклатуры и терминологии [6–24]. В силу бурного роста номенклатурных номинаций, их разнообразия и возрастающей роли в жизни общества, исследование этого языкового материала приобретает все большую актуальность, особенно учитывая, что, как было убедительно показано, помимо чисто научных и промышленно-производственных номенклатурных единиц, существуют и другие номенклатурные обозначения, образующие целые знаковые системы, «живущие» и функционирующие по своим законам [25–27].

В обзорных работах [18, 23] отмечается значительное разнообразие (а порой и несопоставимость) точек зрения по этому вопросу как в философских, так и в лингвистических его рамках, причем намечаются три достаточно длительные общенаучные традиции в употреблении, толковании и понимании термина «номен» и, соответственно, «номенклатура». В учебном пособии [24], допускающем, вообще говоря, весьма различные понимания номенов и номенклатуры, предлагается точка зрения, которая является развитием лингвистического – третьего из перечисленных подходов к пониманию номенклатурных единиц, а для изложения этой точки зрения авторы предлагают собственные, несколько измененные, обозначения. Так, вводится термин «номенклатурное наименование», который совместно с термином «номен» уточняется следующим образом:

1. Номенклатурное наименование – это терминологическое обозначение частного специального понятия какой-либо области знания, дисциплины или

тематической области, которое состоит из двух лексико-синтаксических компонентов, синтаксически главный из которых является термином, словом или словосочетанием общего языка и обозначает специальное родовое понятие данной области, а синтаксически подчиненный – является условным, внешним (для данной области знания и для обозначения соответствующего понятия) знаком, номеном (меткой, «языковой этикеткой») и служит для выделения из родового понятия именно данного частного понятия, фиксируемого в специальных описаниях, толкованиях, дефинициях и т.п. Таковы, например, номинации *первое спряжение, второе спряжение, первое склонение, второе склонение, гипотеза Сепира-Уорфа, закон Иллича-Свитыча-Дыбо, закон Шахматова* (языкознание); *александрійский стих, алкеева строфа, байронический роман, сапфическая строфа* (литературоведение); *тополь белый (Populus alba), клён калифорнийский (Acer negundo), кипарис кашмирский (Cupressus cornuana)* (биология); *пшеница Безостая 4, пшеница Мироновская 808, ячмень Ильинецкий, ячмень Палладиум 596* (сельское хозяйство; примеры из работы [19, с. 8–9]); *лейденская банка, маятник Фуко, эффект Доплера, счётчик Гейгера, твёрдость по Барколу* (физика); *зенитно-ракетный комплекс «Искандер», зенитно-ракетный комплекс «Калибр», зенитно-ракетный комплекс Litner, реактивная система залпового огня «Град», реактивная система залпового огня «Ураган»* (военная техника); *автомобиль ВАЗ 2109, грузовой автомобиль КраЗ-256* (автомобилестроение); *бином Ньютона, алгоритм Герона, многочлены Чебышева, множители Лагранжа, точка разрыва первого рода, гамма-функция (или γ -функция), ряд Тейлора, распределение Гаусса* (математика); *аккумуляторная батарея 6СТ-45 ЭМ, аккумуляторная батарея 6СТ-60 ЭР* (техника) и т.п.¹

¹ Ряд примеров математических и физических номенклатурных наименований взят из словаря [28].

2. Синтаксически подчиненный знак-компонент номенклатурного наименования (метафорически употребленное слово или словосочетание общего языка, антропоним или топоним, аббревиатура или специальная буквенно-цифровая нотация) образует его номенклатурный показатель, его номен (маркер или языковую этикетку). Так, в вышеприведенных примерах «языковым маркером» («языковой этикеткой»), номеном соответственно являются *первое, второе, первое, второе, Сепира-Уорфа, Иллича-Свитыча-Дыбо, Шахматова; александрийский, алкеева, байронический, сапфическая; белый, калифорнийский, кашмирский; Безостая 4, Мироновская 808, Ильинецкий, Палладиум 596; лейденская, Фуко, Доплера, Гейгера, по Барколу; «Искандер», «Калибр», Jitper, «Град», «Ураган»; ВАЗ 2109, КраЗ-256; Ньютона, Герона, Чебышева, Лагранжа, первого рода, гамма, Тейлора, Гаусса; 6СТ-45 ЭМ, 6СТ-60 ЭР.*

3. Понятийное содержание синтаксически главного компонента номенклатурного наименования также может быть результатом метафорических изменений общезыковой лексики, например, *аристотелев фонарь, бритва Оккама, древо Порфирия* (философия), *башия Вячеслава Иванова* (литературоведение), *локон Аньезе* (математика) и т.д.

4. Совокупность всех номенклатурных наименований той или иной области и составляет ее номенклатуру.

Исходя из положений работы [24], в данной статье ставится задача представить классификацию номенов (и содержащих их номенклатурных наименований)². С лингвистической точки зрения изучение номенклатурных наименований чрезвычайно интересно для исследования возможностей языка в деле все более частной, все более дробной номинации различных объектов. Для промышленности и сельского хозяйства способы точной номинации и идентификации сотен тысяч производимых товаров необходимы в деле последующего статистического учета; особенно важно подобное изучение для товароведения, где точная идентификация товара принципиальна во многих отношениях – промышленном, финансовом, юридическом. Поэтому в качестве языкового материала мы обратимся к весьма обширному списку официальных наименований товаров, выпускавшихся и поступавших в торговлю в СССР, – наименований, содержащихся в девятитомном «Товарном словаре» (ТС), изданном в 1956–1961 гг. [30]. Это издание включает описание как продовольственных, так и промышленных товаров, выпускавшихся в СССР; в него включено также описание некоторых видов услуг, оказываемых государственными предприятиями. В связи с тем, что этот словарь имеет уже более чем полувековую историю, приведем его характеристику в виде достаточно пространной цитаты из раздела «От редакции» этого же словаря:

«В Товарном словаре читатель найдет все необходимые сведения о товарах широкого потребления, изготавливаемых отечественной промышленностью и поступающих в продажу населению. Товарный сло-

варь издается в восьми томах³, в которых в общей сложности будет помещено около 8 тыс. статей, описывающих свыше 20 тыс. промышленных и продовольственных товаров широкого потребления. В статьях о продовольственных товарах дается описание всех видов и разновидностей продовольственных товаров, их вкусовых и пищевых свойств, рецептура и химический состав, особенности использования, технология производства, требования, предъявляемые к качеству, и другие данные. В статьях о промышленных товарах описаны промышленные товары широкого потребления всех основных наименований, марок и ряда артикулов. В необходимых случаях приводится назначение товара, применяемые материалы, особенности конструкции, правила пользования, технология производства и др. По всем товарам даются сведения об их потребительских свойствах, правилах хранения, упаковки, маркировки, транспортирования. В создании словаря принимали участие крупные специалисты торговли и промышленности, в числе которых заслуженные деятели наук, доктора и кандидаты наук, профессора и доценты, инженеры и техники, товароведы и другие специалисты» [30, т. 1, раздел «От редакции»] (всюду далее при ссылке на это издание после номера «30» будет использоваться указание тома и столбец в томе, например: 30, т. 2, ст. 246).

Для лучшей ориентации читателя приведем фрагменты двух словарных статей в качестве примеров описания товаров в ТС.

«**Глянцевальный станок** – электрический прибор для сушки фотоотпечатков с одновременным приданием им зеркального блеска. Относится к группе *фототоваров* (см.) Состоит из металлической, слегка изогнутой наружу плиты...» [30, т. 2, ст. 246];

«**Книга** в зависимости от тематики классифицируется Всесоюзной книжной палатой по следующим основным разделам...» [30, т. 3, ст. 635];

«**Хурма быстрозамороженная** – плоды *хурмы субтропической* (см.), замороженные сухим способом (без сахара). Представляет собой ценный пищевой продукт. При замораживании значительно улучшается вкус плодов, исчезает терпкость благодаря разрушению дубильных веществ, содержащихся в свежих плодах и придающих им вяжущий вкус. Используется после дефростации (оттаивания) в качестве десерта. Лучшие для замораживания сорта хурмы...» [30, т. 9, ст. 350].

Заметим, что значительная часть номинаций ТС, называющих те или иные товары и их группы в силу принятых допущений, не представляет собой номенклатурные наименования. В частности, это касается заголовочных единиц (слов или словосочетаний), открывающих так называемые комплексные или групповые словарные статьи. Снова предоставим слово авторам первого тома словаря: «Статьи Товарного словаря... разделяются на три основные группы: 1) комплексные, 2) групповые и 3) так называемые артикульные. В комплексных статьях дается описание совокупности товаров, обладающих общи-

² Применимость данного подхода к лексическому материалу китайского языка изучается в работе [29].

³ Окончательное издание состоит из девяти томов.

ми признаками. Напр., „Вина виноградные“, „Ткани“. В групповых статьях дается характеристика однородных товарных групп или подгрупп. Напр., „Вина крепленые“, „Вина ароматизированные“, „Шелковые ткани“, „Креповые ткани“. Артикульные статьи содержат основные сведения об отдельном товаре. Напр., „Вермут“, „Файдешин“» [30, т. 1, раздел «Как составлен словарь»].

Соответственно, не образуют номенклатурных наименований и не используют в своем составе специальных номенов заголовочные единицы аптекарские товары, духовые музыкальные инструменты, карамельные изделия, кинескоп, клещи, колодки и планки для орденов и медалей, пальтовые ткани, фотоматериалы светочувствительные, хохломские художественные изделия, чернильница конторская стеклянная четырехгранная и т.п. Заголовочные слова так называемых артикульных статей также не являются номенклатурными наименованиями, ср. *кодеин, фуганок, шайба хоккейная, щавель* и т.п. В рамках самых различных словарных статей ТС встречаются весьма протяженные словосочетания, которые идентифицируют производимый товар, но не содержат номенов, ср. *бумажный гофрированный абажур к настольной лампе; бутерброд со шпротами, сардинами и другой рыбой, консервированной в масле; горн кузнечный переносный с вентилятором, приводимым в действие ножным приводом; чернильница конторская с ребристой шторкой; яблочный сок с мякотью* и т.п.; последние являются свободными терминами-словосочетаниями, описательно именуемыми соответствующими торговыми реалиями⁴.

Нас, однако, будут интересовать названия товаров, которые упоминаются и описываются внутри словарных статей ТС и, в силу принятых соглашений, подпадают под рубрику номенклатурных наименований, т.е. наименования типа *автомобиль легковой «Победа М-20», автомобиль легковой «Москвич» 401Б-420; краска алюминиевая светлая АКС-3; эмаль масляная алюминиевая АЭ-8; батарея БС-Г-70, батарея 70-АМЦГ-5 «Дружба»; бритва однозачковая «СТИЗ № 9», велосипед легкодорожный В-32 «Турист»; блесна «Уралка», блесна «Волга»; битумы нефтяные специальные В-100-125°, битумы нефтяные специальные Г-125-135°; бритва электрическая одноножевая Б-53, бритва электрическая двухножевая Б-54; будильник 18Б, будильник 20Б; печенье сахарное «Новая Москва», печенье сахарное «Наша марка», печенье затяжное «Ленинградское»; хлопчатник «108-Ф», хлопчатник «138-Ф»* и т.п.

Каковы же основные типы номенов в составе номенклатурных наименований ТС?

Прежде всего разделим собственно номены номенклатурных наименований на три класса: 1) **вербальные (словесные)**, 2) **буквенно-цифровые** и 3) **смешанные**.

Вербальные номены в своем составе содержат только слова и словосочетания общего языка или термины и не используют последовательностей от-

дельных букв или цифр в качестве самостоятельных компонентов⁵, ср. блесна *«Уралка»*, блесна *«Волга»*; карамель с начинкой *«Гусиные лапки»*, карамель с начинкой *«Кармен»*; картофель *Московский*, картофель *Передовик*, картофель *Ранняя роза*; клей *«Универсал»*, клей синтетический *«Геркулес»*; намордник *«решетка»*, намордник *«уздечка»*; нож промысловый *«Медвежий»*, нож промысловый *«Амурский»*; ночник *«каменный цветок»*, ночник *«ландыш»*, ночник *«теремок»*; орнамент хохломских художественных изделий *«Травка»*, орнамент хохломских художественных изделий *«Круглое письмо»*; печенье сахарное *«Наша марка»*, печенье сахарное *«Новая Москва»* и т.п. Вербальные номены предоставляют богатейший языковой материал для исследования наименований товаров с точки зрения образующих их конструкций, участия в них топонимов, антропонимов и производных от них лексических единиц, с точки зрения мотивированности этих номенов и типов используемых в них метафор, ср. холодильник бытовой *«Саратов»*, холодильник бытовой *«Днепр»*, холодильник бытовой *«Ока»*; шоколад *Басни Крылова*, шоколад *Сказки Пушкина*, шоколад *Москва*, шоколад *«Красный Октябрь»*; часы наручные мужские *«Сатурн»*, часы наручные женские *«Эра»*; карамель леденцовая *«Прозрачная»*, карамель леденцовая *«Спорт»*, монпансье леденцовое с защитным слоем *«Цветной горошек»*.

Буквенно-цифровые номены, в противоположность вербальным, не используют каких-либо общеязыковых единиц или терминов, а используют исключительно последовательности букв и цифр, не образующих слов или терминов, ср. упомянутые выше номенклатурные наименования *автомобиль повышенной проходимости ГАЗ-63; краска алюминиевая светлая АКС-3, эмаль масляная алюминиевая АЭ-8; батарея БС-Г-70; битумы нефтяные специальные Г-125-135°; будильник 20Б; хлопчатник «108-Ф»*, а также номенклатурные наименования *витамин В₁, витамин В₆, витамин К₃; клей синтетический БФ-3, клей синтетический БФ-5; фонарь электрический карманный ФК, фонарь электрический карманный ФК-1; часы маятниковые пружинные настенные М-9, часы маятниковые пружинные настенные М-12; электрофритюрница настольная ЭФ-2, электрофритюрница настольная ЭФ-10*.

Наконец, третий тип составляют номены, использующие в своем составе и нарицательные или собственные имена, и буквенно-цифровые обозначения и имеющие в силу этого смешанный характер, ср. *автомобиль легковой «Победа М-20», автомобиль легковой «Москвич» 401Б-420; батарея 70-АМЦГ-5 «Дружба», батарея 70-АМЦГ-5-п «Энергия»; объектив фотографический «Мир-1», объектив фотографический «Юпитер-11»; палатка туристская «Памирка-2»; перец стручковый острый Астраханский-628, перец стручковый сладкий Болгар*

⁴ По нашим наблюдениям, ни одно заголовочное слово ТС не является номенклатурным наименованием; соответственно, и его словарные статьи всегда начинаются с описания некоторого объекта и связанного с ним понятия.

⁵ Кроме специальных слов-идентификаторов (подобных словам «марка», «модель», «артикул», «тип» и др.), о которых говорится ниже.

*ский-35; среднеформатный фотоаппарат «Любитель-2» и т.п.*⁶

В пределах номенклатурных наименований одной и той же группы товаров могут использоваться номены разных типов, ср. *котлы отопительные «Луч»* и *котлы отопительные «МГ-2»*, *котлы отопительные «Универсал-3»*; *магнитофон «Комета»* и *магнитофон «Днепр-11»*, *магнитофон «МЭЗ-15»*; *фотоаппарат «Зоркий»* и *фотоаппарат «Зоркий-2С»*, *фотоаппарат «Мир»* и *фотоаппарат «Киев-ША»*; *ячмень двурядный «Винер»*, *ячмень многорядный «Красный дар»* и *ячмень двурядный «Прекочиус 143»*.

В этих условиях закономерен вопрос, важный и для лингвистических, и для производственно-технологических, и для коммерческо-юридических аспектов обозначения товара: чем объясняется использование разных типов номенов при обозначении одной и той же группы товаров? Объясняется ли смена типа номена изменением технических параметров товара, технологии его производства и т.п., или она может быть совершенно произвольной? И, наконец, самый важный вопрос номинации для производственной сферы: должна ли смена номена как-то ориентироваться на общность с предшествующими его производственно-технологическими параметрами, показателями и, если да, то как? Имеется ли какая-либо нормативная база, фиксирующая решение подобных вопросов, и должна ли она меняться и как часто? Или, наконец, она только еще должна быть разработана?

Говоря о принятом классификационном членении номенов, следует отметить, что в некоторых случаях их дальнейшее классификационное деление представляется очевидным. Так, буквенно-цифровые номены легко подразделяются на чисто буквенные (т.е. состоящие только из букв), чисто цифровые и собственно буквенно-цифровые.

Основную часть номенов этого типа составляют буквенно-цифровые номены (включающие в свой состав и буквы, и цифры), ср. приведенные выше наименования *автомобиль повышенной проходимости ГАЗ-63*; *краска алюминиевая светлая АКС-3*; *эмаль масляная алюминиевая АЭ-8*; *батарея БС-Г-70*; *битумы нефтяные специальные Г-125-135°*; *клей синтетический БФ-3*, *клей синтетический БФ-5*; *фонарь электрический карманный ФК-1*; *часы маятниковые пружинные настенные М-9*; *часы маятниковые пружинные настенные М-12*; *электрофритюрница настольная ЭФ-1*, *электрофритюрница настольная ЭФ-2*.

Чисто буквенных номенов в составе номенклатурных наименований ТС сравнительно немного, таковы: *автомобильная карданная смазка АМ*; *вита-*

мин А, *витамин С*, *витамин РР*; *индустриальная смазка для подшипников ИПК (смазка Каретникова)*, *смазка текстильная ИТ*; *фонарь электрический ФК*; *фотоаппарат малоформатный ФЭД* и др.

Чисто цифровые номены, состоящие только из цифр и «не отягощенные» какими-либо иными компонентами, по нашим наблюдениям, в исследуемом материале отсутствуют.

Здесь, однако, необходимо обратить внимание на использование при формулировке номенклатурных наименований в отдельных случаях особых словесных идентификаторов, к которым применительно к анализируемому материалу, с нашей точки зрения, следует отнести единицы «*артикул*», «*класс*», «*марка*», «*модель*», «*модификация*», «*образец*», «*порода*», «*сорт*», «*тип*» и «*N*». Перечисленные идентификаторы вводят те или иные номены и без них не употребляются, ср. невозможность самостоятельного использования *сапоги резиновые клееные артикула**, *рукава напорные типа**, *очки предохранительные N**, *репчатый лук сорта** и т.п.

Относительно артикула в ТС сообщается: «**Артикул** – условное обозначение, присваиваемое товару в целях отличия его от другого, аналогичного ему товара по несущественным признакам. Для обозначения артикула применяются цифры, различные названия и имена, отдельные буквы и сочетания букв (не дающие слов), цифры с буквами, названия с цифрами и другие комбинации» [30, т. 1, с. 230]. Известно, что введение артикулов заменяет подробное описание товаров особым обозначением, тем самым ускоряя ведение торговой документации, учета товаров, облегчает изучение спроса и составления заказов на поставки товаров. При использовании слов-идентификаторов «чисто» цифровые номены вполне возможны: так, в исследованном материале с помощью артикула (в сокращенной форме – «*арт.*») обычно характеризуются товары текстильной, обувной, меховой промышленности, ср. *галошная ткань арт. 4250*, *детская матроска арт. 2147*, *занавесочная ткань арт. 1272*, *капроновая ткань блузочная арт. 3255*, *коверкот меланжевый арт. 748*, *сапоги резиновые клееные арт. 150*, *трико костюмное арт. 703*, *шинельное сукно арт. 4412*, *штоф арт. 1407*. В ряде случаев идентификатор «артикул» добавляется к уже имеющемуся вербальному номену, в результате чего выстраивается определенная иерархия номенов: *муслин блузочный Незабудка арт. 32140*, *креп Мария арт. 31129*, *креп Лилия арт. 3192*.

Слова-идентификаторы «сорт» и «порода» используются в разных тематических областях: первые – для номенклатурных обозначений группы однородных товаров в растениеводстве и пищевой промышленности, вторые – в животноводстве (а также для обозначения древесины, получаемой из деревьев различных родов и видов в деревообрабатывающей промышленности); ср. *репчатый лук сорта Арзамасский*, *репчатый лук сорта Барлетта*, *репчатый лук сорта Стригуновский*; *цикорий сорта Борисовский (исполинский)*, *цикорий сорта Силезский* и т.д.

Номенклатурный идентификатор «модель» часто используется в обозначении марок транспортных

⁶ Использование вербальных и смешанных номенов ставит разнообразные вопросы их лингвистического описания – с точки зрения используемой синтаксической конструкции, входящих в состав номена антропонимических или топонимических единиц, вида используемой метафоры и т.п. В настоящее время имеются лишь отдельные немногочисленные фрагментарные описания этого огромного и разнообразнейшего языкового материала (даже если иметь в виду только товарно-производственную номенклатуру), ср. работы [15, 16]. В настоящей работе мы, разумеется, не имеем возможности останавливаться на этом вопросе.

средств, ср. *мотоцикл легкий модели К-175*, *мотоцикл легкий модели М1А*; *автомобиль грузовой модели ГАЗ-51*.

Номены с идентификаторами «типа», «класса» или «N» могут быть в составе номенклатурного обозначения товара, использующего какие-либо вербальные компоненты (ср. *минеральная вода эссенцуки N 4*, *минеральная вода эссенцуки N 17*; *молоток отбойный русского образца*) или не использующего таковых. Последние представляют особый интерес: ср. *микроскоп детский типа МУ*; *пододеяльник типа А*, *пододеяльник типа Г*; *спринцовки резиновые типа А*, *спринцовки резиновые типа Б*; *шнур резиновый типа II*, *шнур резиновый типа III*; *рукава напорные типа Б*, *рукава напорные типа В*, *рукава напорные типа II*, *рукава напорные типа III*; *электропила плотничная типа И-7*; *яхты спортивные килевые типа Л-4*; *швертбот класса Р-24*; *лак мебельный нитроглифталевый N 7554*; *лак мебельный нитроглифталевый N 757*, *нити для штопки N 25*, *нити для штопки N 40*; *хомут для сельскохозяйственной и тяжеловозной упряжи тяжелый N 1*, *хомут для сельскохозяйственной и тяжеловозной упряжи тяжелый N 2* и т.п. Подобные номены вполне аналогичны чисто буквенным или чисто цифровым, ср. потенциально вполне возможные, но реально не представленные *микроскоп детский МУ**; *пододеяльник А**, *пододеяльник Г**; *спринцовки резиновые А**, *спринцовки резиновые Б**; *шнур резиновый II**, *шнур резиновый III**; *рукава напорные Б**, *рукава напорные В**, *рукава напорные II**, *рукава напорные III**; *швертбот Р-2**. Совершенно аналогичное можно утверждать для номенклатурных наименований ТС, содержащих номены «N 1», «N 2», «N 3» и т.п., например: *лак мебельный нитроглифталевый 7554**; *нити для штопки 25**, *нити для штопки 40** и т.д.

Заметим, что в случае использования различных идентификаторов описание товара осуществляется в целом так же, как и в случаях использования номенов без идентификаторов, например:

«**Микроскоп...** Биологический микроскоп *типа МБИ-4* является микроскопом дорожного типа; приспособлен для работы в экспедиционных условиях. Отличие его от *МБИ-1* заключается ...» [30, т. 6, ст. 682];

«**Пододеяльники** выпускаются четырех типов – *А, Б, В и Г...* Тип *А* (рис. 1) – открытые, прямоугольной формы, односторонние, с бортиками на петлях...» [30, т. 7, ст. 160];

«**Шпаклевки перхлорвиниловые N 175 и N 176** – густотертые пасты на глифталевой основе. Применяются при окраске автомобилей в качестве промежуточного покрытия по грунту N 138 для выравнивания загрунтованной поверхности под нитроэмали...» [30, т. 9, ст. 746];

«**Модель 21И** отличается несколько иной облицовкой радиатора, бамперами, подфарниками, радиоприемником; укомплектован опрыскивателем лобового стекла» [30, т. 9, ст. 115] (речь идет о легковом автомобиле «Волга»);

«**Швертбот класса Р-2**. Размеры (в м.): длина не ограничена, наибольшая шир. 1,7; выс. борта 0,30;

вес шверта 80 кг; парусность 20 м²; с полупалубой. Предназначается для...» [30, т. 9, ст. 1215].

Представляется, что задача номенов с идентификаторами в составе товарных номенклатурных наименований та же, что и номенов без идентификаторов – достижение более точной, более частной номинации объекта и сужение (часто по многим признакам одновременно) родового понятия, из которого этот объект вычленяется.

Для лингвистической характеристики номенов не менее существенно их деление по числу и разнообразию входящих в них самостоятельных отдельных последовательностей – будь то словесные, буквенные или цифровые последовательности. В этой связи будем считать номен **простым**, если его формулировка состоит только из одного слова или из одной отдельной последовательности букв (не являющейся словом) или цифр⁷; если же номен содержит две, три или более указанных последовательности будем считать его **сложным**. Соответственно, номены наименований блесна «*Уралка*», блесна «*Волга*»; будильник *17Б*, будильник *20Б*; нож складной *У9*, нож складной *У10А*; холодильник бытовой «*Днепр*», холодильник бытовой «*Ока*»; часы карманные «*Луч*», часы наручные женские «*Эра*»; карамель леденцовая «*Прозрачная*», карамель леденцовая «*Спорт*» – это простые номены; все смешанные номены, уже в силу определения, являются сложными, их примеры см. выше. Сложными являются также номены наименований *батарея БС-Г-70*; *битумы нефтяные специальные Г-125-135°*; *клей синтетический БФ-2*, *клей синтетический БФ-3*, *клей синтетический БФ-5*; *краска алюминиевая светлая АКС-3*; *настойка «Нежинская рябина»*, *настойка «Рижский черный бальзам»*; *патиссоны «Белый 13»*, *патиссоны «Белый тарельчатый»*; *персики «Майский ветерок»*, *персики «Золотой юбилей»*; *пианино «Красный Октябрь»*; *фонарь электрический карманный ФК-1*; *хлопчатник «108-Ф»*, *хлопчатник «138-Ф»*; часы маятниковые пружинные настенные *М-9*; часы маятниковые пружинные настенные *М-12*; *эмаль масляная алюминиевая АЭ-8*.

Вопрос о том, в какой степени употребление единиц «марка», «модель», «модификация», «образец», и т.д. обусловлено лишь традицией, а в какой – жестко предписывается нормативной и производственно-технологической документацией, подлежит специальному исследованию, и его решение не входит в задачи настоящей статьи. Вполне возможно, что техническая документация в ряде случаев предусматривает использование лишь одной единицы из этого ряда, а в общелитературном языке они сближаются. Следует все же отметить, что в «Новом политехническом словаре» одно из значений слова «модель» толкуется так: «Образец, служащий эталоном для серийного или массового воспроизведения какого-либо изделия..., а также тип, марка изделия или конструкции» [31, с. 305]; аналогично синонимизируются в

⁷ Последовательность букв или цифр естественно считать отдельной, если внутри она не разделяется на письме пробелом или знаком препинания.

соответствующих значениях слова «модель», «тип», «марка» и в энциклопедии [32, с. 966].

Кратко остановимся на вопросах правильного написания номенклатурных наименований и входящих в них номенов, совместно образующих названия товаров и продукции. Эти вопросы включают целый ряд пунктов, и в первую очередь они касаются использования кавычек и прописной или строчной первой буквы слова (или словосочетания) в кавычках. Следует заметить, что на момент опубликования ТС действовали «Правила русской орфографии и пунктуации» (М.: Учпедгиз, 1956), в которых раздела о написании названий товарных марок и сортов не было. В этой связи заметим, что в рассмотренном материале написание номенклатурных наименований и номенов далеко не всегда последовательно, ср. *мотор лодочный подвесной «Москва»* (в кавычках), но *мотор лодочный подвесной Рига-125* (без кавычек); *стиральный порошок «Новость»*, *стиральный порошок «Кристалл»* (в кавычках), но *мыло туалетное Красная Москва*, *мыло туалетное Московское* (без кавычек); *автомобиль легковой «ЗИМ»*, но *автомобиль легковой ЗИС-150* (без кавычек) или *ложка кухонная «уполовник»*, *батон городской*.

В настоящее время орфографическое и пунктуационное оформление номенклатурных названий товаров и продукции подчиняется ряду документов – справочников и правил, ср. [33–35]. В то же время как в пределах одного документа, так и – особенно – при их сравнении обнаруживаются некоторые неувязки и неясности.

В справочнике Д.Э. Розенталя и др. [33], как представляется, приведены не вполне точные правила написания видов и сортов растений. Так, в нем сообщается: «Кавычками выделяются названия сельскохозяйственных культур, овощей, цветов и т.д., например: пшеница «крымка», клубника «виктория», гладиолус «элегия»» [33, с. 26]. В другом месте того же справочника утверждается: «Названия видов и сортов растений, овощей, фруктов, цветов в специальной литературе (выделено нами – С.Ш.) пишутся с прописной буквы, например: малина Мальборо, земляника Победитель, смородина Выставочная красная, слива Никольская белая, морковь Нантская ... тюльпан Черный принц» [33, с. 25]. Наконец, там же, по-видимому, также по отношению к специальной литературе представлена еще одна рекомендация: «В текстах, не перегруженных названиями сортов растений, овощей, фруктов и т.д., эти названия заключаются в кавычки и пишутся со строчной буквы, например: клубника «виктория», помидор «иосиф прекрасный»» [33, с. 25; выделено нами – С.Ш.]. Таким образом, имеются две возможности написания видов и сортов культурных растений в зависимости от условий употреблений номенклатурного наименования. Так, в специальной литературе пишут земляника Победитель, смородина Выставочная красная, слива Никольская белая, морковь Нантская, картофель Эпикур, пшеница Днепровская-521, фиалка Пармская, тюльпан Черный принц, а в неспециальной, по-видимому, – земляника «победитель», смородина «выставочная красная», тюльпан «черный принц». Наконец, в текстах специальной литературы,

«не перегруженных названиями сортов растений, овощей, фруктов» допускается написание, фактически совпадающее с написанием в неспециальной литературе: земляника «победитель», смородина «выставочная красная», тюльпан «черный принц» и помидор «иосиф прекрасный». Заметим также, что выражения «специальная литература», «тексты, не перегруженные названиями сортов растений» никак не поясняются и, безусловно, могут быть источником разночтений.

Аналогичны рекомендации справочника [34, с. 103], хотя «тексты, не перегруженные названиями сортов растений, овощей, фруктов» здесь не упоминаются. Подобные же требования к написанию номенклатурных наименований содержатся и в своде правил [34, с. 188]: «Названия видов и сортов сельскохозяйственных культур, овощей, цветов и т. п. – термины агрономии и садоводства – выделяются кавычками и пишутся со строчной буквы, напр.: пшеница «крымка», клубника «виктория», малина «Мальборо», тюльпан «чёрный принц», слива «Никольская»». Здесь, однако, не упоминаются ни «специальная литература», ни «тексты, не перегруженные названиями сортов растений, овощей, фруктов».

Имеются некоторые разногласия и в рекомендации написания названий машин, самолетов и т.п. Так, хотя рекомендации писать названия марки машин в работах [33–35] в целом совпадают (их следует писать в кавычках и с прописной буквы – автомобиль «Волга», автомобиль «Жигули», автомобиль «Чайка»; самолет «Руслан», самолет «Фантом», самолет «Боинг-707», самолет «Мессершмитт-109», самолет «Хейнкель-111»), справочник Розенталя допускает при передаче разговорной речи возможность написания *танки «тигр»*, *танки «леопард»* [33, с. 197], а «Справочник издателя и автора» такой возможности не оговаривает.

Подведем теперь некоторые итоги.

При идентификации товаров, помимо общезыковых развернутых описательных конструкций русского языка, используются различные номены, которыми обозначаются группы различных, но (на определенном уровне иерархии) однородных товаров. Сами номены могут быть буквенно-цифровыми, вербальными и смешанными. В составе номенов могут присутствовать или не присутствовать идентификаторы из весьма ограниченного списка типа «артикул», «марка», «модель», «модификация», «образец» и др. с добавлением соответствующих вербальных, буквенных или цифровых последовательностей. Путем применения нескольких идентификаторов, использующих те или иные номены, создается иерархия уточнения при номинации того или иного товара. Описание общности номенов по одному или нескольким важнейшим для номенов критериям задает те или иные их языковые типы, после чего можно поставить вопросы о зависимости различных типов номенов от различных внемаркетинговых факторов номинации и последовательного использования соответствующего типа номена в производственной практике, об унификации номенов для одной или нескольких различных групп товаров и т.п.

Остаются актуальными и вопросы орфографии номенклатурных наименований, касающиеся исполь-

зования или неиспользования кавычек, оформляющих номены, а также их написания со строчной или прописной буквы. Настоящая работа, как можно надеяться, могла бы способствовать более точной постановке этих вопросов, их практическому решению и внедрению этих решений в промышленное и сельскохозяйственное производство.

СПИСОК ЛИТЕРАТУРЫ

1. Уэвель В. История индуктивных наук от древнейшего и до настоящего времени. Т. 3. – СПб.: Тип. Н. Тиблена и комп., 1869. – 930 с.
2. Милль Д.С. Система логики силлогической и индуктивной. Изложение принципов доказательства в связи с методами научного исследования / пер. с 10-го англ. изд. под ред. В.И. Ивановского. – М.: Изд-е магазина «Книжное дело», 1899. – 782 с.
3. Винокур Г.О. О некоторых явлениях словообразования в русской технической терминологии // Труды Моск. ин-та истории, философии и литературы. Т. V. Сб. статей по языковедению. – М., 1939. – С. 3–54.
4. Реформатский А.А. Термин как член лексической системы языка // Проблемы структурной лингвистики, 1967. – М.: Наука, 1968. – С. 103–125.
5. Реформатский А.А. Что такое термин и терминология. – М., 1959. – 14 с.
6. Ахманова О.С. Словарь лингвистических терминов. – М.: Сов. энцикл., 1966. – 608 с.
7. Шарашова М.К. К разграничению терминов и номенклатурных единиц (на примере названий растений) // Уч. зап. Калининского пед. ин-та им. М.И. Калинина. – 1969. – Т. 66, Ч. 1. – С. 75–81.
8. Хаяутин А.Д. Термин, терминология, номенклатура: учеб. пособие. – Самарканд: Самарканд. гос. ун-т им. А. Навои, 1972. – 129 с.
9. Канделаки Т.Л. К вопросу о номенклатурных наименованиях // Вопросы разработки научно-технической терминологии. – Рига, 1973. – С. 60–70.
10. Рейцак А. Терминология и номенклатура // Вопросы разработки научно-технической терминологии. – Рига, 1973. – С. 71–91.
11. Лейчик В.М. Номенклатура – промежуточное звено между терминами и собственными именами // Вопросы терминологии и лингвистической статистики. – Воронеж, 1974. – С. 13–24.
12. Березникова Р.Е. Место номенов в лексической системе языка // Имя нарицательное и собственное. – М.: Наука, 1976. – С. 42–58(а).
13. Березникова Р.Е. Подача номенов в словарях различных типов // Проблематика определений терминов в словарях разных типов. – Л.: Наука, 1976. – С. 83–91(б).
14. Забинкова Н.Н. Термины и номенклатурные слова в ботанических словарях // Проблематика определений терминов в словарях разных типов. – Л.: Наука, 1976. – С. 91–101.
15. Комарова З.И. Выявление мотивировочного признака фитонимов (на материале номенклатурных названий сортов картофеля) // Проблемы ономазиологии. IV. (Науч. тр. Курск. пед. ин-та. Т. 175). – Курск, 1976. – С. 70–73.
16. Комарова З.И. О некоторых закономерностях «искусственной» номинации (на материале названий сортов картофеля) // Вопросы ономастики. Вып. 13. – Свердловск, 1979. – С. 116–121.
17. Суперанская А.В. Терминология и номенклатура // Проблематика определений терминов в словарях разных типов. – Л.: Наука, 1976. – С. 73–83.
18. Шелов С.Д. Об одном классе научно-технической лексики (три подхода к выделению номенклатурных наименований) // Научно-техническая информация. Сер. 2. – 1985. – № 3. – С. 1–7.
19. Комарова З.И. Семантическая структура специального слова и ее лексикографическое описание. – Свердловск: Изд-во УрГУ, 1991. – 156 с.
20. Суперанская А.В., Васильева Н.В., Подольская Н.В. Общая терминология. Терминологическая деятельность. – М.: Ин-т этнологии и антропологии РАН, 1993. – 288 с.
21. Оленева Т.Б. Термины и номенклатура в системе диалектной лексики // Вопросы терминологии: межвуз. сб. – Нижний Новгород, 1993. – С. 52–65.
22. Лейчик В.М. Создание и использование номенклатурных единиц в современном русском языке // Стереотипность и творчество в тексте. Вып. 8: межвуз. сб. науч. тр. / под ред. М.П. Котуровой. – Пермь, 2005. – С. 156–162.
23. Шелов С.Д. Номенклатура и терминология (три подхода к выделению номенклатурных наименований) // Актуальные проблемы лингвистики и терминоведения. – Екатеринбург, 2007. – С. 76–82.
24. Шелов С.Д., Лейчик В.М. Номенклатурные наименования как класс научно-технической лексики: состав и функции: учеб. пособие. – СПб., 2007. – 80 с.
25. Лазарева М.Н. Когнитивный аспект структурно-семантической организации фитонимов, функционирующих в прикладной науке (на материале французского языка) // Терминология и знание. Материалы III Междунар. симп. – М., 2013. – С. 283–294.
26. Лазарева М.Н. Универсальное и лингвоспецифичное в научной картине мира (на примере латинской и русской ботанических номенклатур) // Терминология и знание. Вып. IV. Материалы IV Междунар. симп. – М., 2014. – С. 101–109.
27. Бурдина О.Б. Мишланова С.Л. Особенности терминологического моделирования в дискурсе смешанного типа (на материале номинации лекарственных препаратов) // Терминология и знание. Вып. IV. Материалы IV Междунар. симп. – М., 2014. – С. 70–83.

28. Галевский Г.В., Мауэр Л.В., Жуковский Н.С. Словарь по науке и технике (Английский. Немецкий. Русский). – М.: Флинта; Наука, 2003. – 320 с.
29. Сунь Шуфан, Шелов С.Д. Номенклатурные наименования как элемент китайской лексики (на материале языкознания и литературоведения) // Вестник Санкт-Петербургского университета. Сер. Востоковедение. Африканистика. – Вып. 3. – 2014. – С. 5–16.
30. Товарный словарь / гл. ред. И.А. Пугачев. Т. 1–9. – М.: Гос. изд-во торговой литературы, 1956–1961.
31. Новый политехнический словарь / Гл. ред. А.Ю. Ишлинский. – М.: Большая Российская энциклопедия, 2000. – 671 с.
32. Большая Российская энциклопедия. – М.: Большая Российская энциклопедия, 2005. – 1888 с.
33. Розенталь Д.Э., Джанджакова Е.В., Кабанова Н.П. Справочник по русскому языку. Правописание, произношение, литературное редактирование. 7-е изд. – М.: Айрис-пресс, 2010. – 496 с.
34. Мильчин А.Э. Справочник издателя и автора. Редакционно-издательское оформление издания. 3-е изд. – М.: Изд-во Студии Артемия Лебедева, 2009. – 1084 с.
35. Правила русской орфографии и пунктуации. Полный академический справочник / под ред. В.В. Лопатина. – М.: Эксмо, 2006. – 480 с.

Материал поступил в редакцию 17.02.15.

Сведения об авторах

Е Цисун – кандидат филологических наук, доцент, Центр изучения русской филологии и культуры Хэйлунцзянского университета Министерства образования КНР, г. Харбин.
e-mail: Yeqisong@163.com

ШЕЛОВ Сергей Дмитриевич – доктор филологических наук, ведущий научный сотрудник Института русского языка РАН, Москва.
e-mail: Volehs@mail.ru